

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2005, 05:50:41 ; Search time 4631 Seconds
(without alignments)
10609.715 Million cell updates/sec

Title: US-09-842-628-1
Perfect score: 1014
Sequence: 1 aaggaagtggaagaccagca.....ttctcgggaactactgataa 1014

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: gb Da:*
2: gb Hcg:*
3: gb In:*
4: gb Com:*
5: gb Ov:*
6: gb Pat:*
7: gb Ph:*
8: gb Pl:*
9: gb Pr:*
10: gb Ro:*
11: gb Sts:*
12: gb Sy:*
13: gb Un:*
14: gb Vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	983.4	97.0	1017	6 AR177619	AR177619 Sequence
2	823.6	81.2	308015	1 AE016783	AE016783 Pseudomon
3	740.4	73.0	9937	1 AE004563	AE004563 Pseudomon
4	738.8	72.9	1089	12 AY659070	AY659070 Synthetic
5	707.8	69.8	1709	1 AF056495	AF056495 Pseudomon
6	610.4	60.2	2353	1 PFL238710	PFL238710 Pseudomon
7	313	30.9	110000	1 CR543861.20	CR543861.20 Continuation (21 o
8	283.2	27.9	110000	1 BX571965.16	BX571965.16 Continuation (17 o
9	281.6	27.8	110000	1 CR000010_14	CR000010_14 Continuation (15 o
10	281	27.7	300	6 A27539	A27539 Pseudomonas
11	273	26.9	1125	6 AR319786	AR319786 Sequence
12	266.4	26.3	1399	1 ECASN	ECASN
13	265.4	26.2	1047	1 AY560098	AY560098 Erwinia chr
14	264.6	26.1	348251	1 BX640423	BX640423 Bordetella
15	264	26.0	347356	1 BX640437	BX640437 Bordetella
16	261.6	25.8	2450	1 ECSC	ECSC
17	261.6	25.8	2837	6 A14577	A14577 Erwinia chr
18	261.6	25.8	2837	6 B01113	B01113 Genomic DNA
19	260.8	25.7	346359	1 BX640411	BX640411 Bordetella

C	20	255	25.1	10893	1	AE015308	AE015308 Shigella
C	21	255	25.1	292906	1	AE016988	AE016988 Shigella
C	22	250.4	24.7	110000	1	AP006841.48	Continuation (49 o
C	23	250.2	24.7	10749	1	AE005526	AE005526 Escherich
C	24	250.2	24.7	266658	1	AP002563	AP002563 Escherich
C	25	245.4	24.2	303121	1	AE016766	AE016766 Escherich
C	26	241.8	23.8	1530	1	ECOLASNT	M34277 E. coli L-as
C	27	241.8	23.8	1643	1	ECOLASNT	M34277 E. coli L-as
C	28	241.8	23.8	1643	1	ECOLASNT	M34277 E. coli L-as
C	29	241.8	23.8	14744	1	U00096.30	Continuation (31 o
C	30	240.8	23.7	1848	6	ECU28377	U78377 Escherichia
C	31	240.8	23.7	1848	12	I73512	I73512 Sequence 15
C	32	240.4	23.7	1071	6	AR384001	U06943 Synthetic c
C	33	230.8	22.8	305961	1	AR016937	AR016937 Sequence
C	34	230.4	22.7	304500	1	AP005953	AP005953 Bradyrhiz
C	35	228.8	22.6	303450	1	AP005085	AP005085 Vibrio pa
C	36	228.4	22.5	1041	1	AY560097	AY560097 Pectobact
C	37	228.4	22.5	1177	1	AY327441	AY327441 Pectobact
C	38	228.4	22.5	110000	1	BX50851.12	Continuation (13 o
C	39	219	21.6	1035	1	MSANSAGEN	X83689 W. succinoge
C	40	219	21.6	1133	6	AR157944	AR157944 Sequence
C	41	219	21.6	1133	6	AX464427	AX464427 Sequence
C	42	219	21.6	2505	1	MSDCUANGA	X89215 W. succinoge
C	43	219	21.6	346792	1	BX571658	BX571658 Wolinella
C	44	213	21.0	110000	1	BX571965.27	Continuation (28 o
C	45	209	20.6	207050	1	AL646063	AL646063 Ralstonia

ALIGNMENTS

RESULT 1	AR177619	Sequence 1 from patent US 6312939.	1017 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	AR177619					
DEFINITION	Sequence 1 from patent US 6312939.					
ACCESSION	AR177619					
VERSION	AR177619.1	GI:17919974				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	Unclassified.					
AUTHORS	1 (bases 1 to 1017)					
TITLE	Roberts,J., MacCallister,T.W., Sethuraman,N. and Freeman,A.G.					
JOURNAL	Genetically engineered glutaminase and its use in antiviral and anticancer therapy					
FEATURES	Patent: US 6312939-A 1 06-NOV-2001;					
source	Location/Qualifiers					
	1..1017					
	/organism="unknown"					
	/mol_type="unassigned DNA"					

ORIGIN

Query Match	97.0%; Score 983.4; DB 6; Length 1017;
Best Local Similarity	98.6%; Pred. No. 3e-108;
Matches 1003; Conservative 0; Mismatches 11; Indels 3; Gaps 1;	
QY	1 AAGGAATGAGAAACGACGAGCTGCGCAAGTGTATCTTGGCCACCGGGGACCC 60
DB	1 AAGGAATGAGAAACGACGAGCTGCGCAAGTGTATCTTGGCCACCGGGGACCC 60
QY	61 ATGCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB	61 ATGCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY	121 GTCGACAGCTGATATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 180
DB	121 GTCGACAGCTGATATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 180
QY	181 CAGGTATGCAATCGCTCCGAAAGCATCAACGACGACCTGCTCAAGCTGGCAAGC 240
DB	181 CAGGTATGCAATCGCTCCGAAAGCATCAACGACGACCTGCTCAAGCTGGGCAAG 240
QY	241 AGCGTGGCGAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 300

Db 916 TCGACGCTCAATGCGCGGCTTCCTGCGCAACGCCGAGCAGCCGACGACGACGAC 975

Qy 898 GACTGGGTGCTGCGCCACGACCTGAAACCCGAGAAAGCCCGCATCTGCGGATGTGCA 957

Db 976 GACTGGATGCTGCGCCACGACCTGAAACCCGAGAAAGCGCGCATCTGCGGATGTGCA 1035

Qy 958 ATGACCAAGACCCGAGCAGCAGCAGAGAGCTGACGCGCATTTTCTGGGAATACTGA 1011

Db 1036 ATGACCAAGACCCGAGCAGCAGCAGAGAGCTGACGCGCATTTTCTGGGAATACTGA 1089

RESULT 5
AF056495 1709 bp DNA linear BCT 08-OCT-1999
LOCUS Pseudomonas fluorescens glutaminase-asparaginase precursor, gene,
DEFINITION complete cds.
ACCESSION AF056495
VERSION AF056495.1 GI:3044186
KEYWORDS Pseudomonas fluorescens
SOURCE Pseudomonas fluorescens
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1709)
Huser, A., Klopner, U. and Rohm, K.H.
AUTHORS Cloning, sequence analysis, and expression of anah from Pseudomonas
TITLE fluorescens, encoding periplasmic glutaminase/asparaginase
JOURNAL FEMS Microbiol. Lett. 178 (2), 327-335 (1999)
MEDLINE 99429098
PUBMED 10499283
REFERENCE 2 (bases 1 to 1709)
Rohm, K.H. and Huser, A.
AUTHORS Direct Submission
TITLE Submitted (01-APR-1998) Physiol. Chemie, Uni-Marburg, Karl von
JOURNAL Fritsch Str.1, Marburg/ Lahn D - 35033, Germany
FEATURES
source
1. 1709
/organism="Pseudomonas fluorescens"
/mol_type="genomic DNA"
/strain="ATCC 13525"
/db_xref="ATCC:13525"
/db_xref="taxon:294"
295. 1383
/BC_number="3.5.1.38"
/note="periplasmatic"
/codon_start=1
/transl_table=1
/product="glutaminase-asparaginase precursor"
/protein_id="AAC33155.1"
/db_xref="GI:3044187"
/translation="MRSALKTPVPGALALLLPVNAQAKVETKTGLANVVIATGSG
TIAAGASANSATYQAKVIGIOLIAVGBLSGINAVRSGQVQIASISINNENLQ
LGRVAVLADSKVDGIIVTHGDTLBEYFLNLVKTDPKPIIVGSMRPTMSAD
GMLNLYNAVAVAGSKDARAGKGLVLTNMBELIOSGADVSKMINIKTEAKSPWGLAMTV
EGKSYFRLPAKRTMDSBPDIKTKSLPVEIAYGYGNSDTLVKTLAAGAAIITH
AGTNGSVSKVVPALQELRKQGVQIIRSHVNAAGFVLNNAEQPDKIKWVVAHDLN
PQAKITAMVALTTQDSKELQRMFMY"
295. 369
mat_peptide
370. 1380
/product="glutaminase-asparaginase"

ORIGIN

Query Match 69.8%; Score 707.8; DB 1; Length 1709;
Best Local Similarity 87.8%; Pred. No. 1.8e-75;
Matches 830; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

Qy 1 AAGGAATGAGAGCAGCAGCAGAGCTGGCCAAAGTGGATCTGCGCAACCGCGGAGACC 60

Db 370 AAGGAATGAGAGCAGCAGCAGAGCTGGCCAAAGTGGATCTGCGCAACCGCGGAGACC 429

Qy 61 ATGCGCGGCGTGGCGCCAGCGCGCCCAACAGCGCCACTTACAGAGCTGCCAAGGTTGCG 120

Db 430 ATGCGCGGCGCGCGCGCCAGCGCGCCCAACAGCGCCACTTACAGAGCGCGCCAAGTTCGGT 489

Qy 121 GTGACAAGCTGATTTGCGCGGCTGCGGAGCTGGCCGACCTGGGCAATGTGCGGCGAG 180

Db 490 ATGAGCAATTTGATGCGCGGCTGCTGAGCTTTAGCCAGATGCGCAATGTCGCGGAG 549

Qy 181 CAGGTGATGAGATGCGCTCCGAAGAATCAACCAAGACGACTTCTCAAGCTGGCAAGC 240

Db 550 CAAGTATGAGAAATGGGTCCGAGAGCATCAACAAAGAAACCTTCTGCAATGGGTGCG 609

Qy 241 AGCGTGGCCGAGCTGGCGCAGACGAAATGACGTGATGGCATGCTCATCAACCAAGCAC 300

Db 610 CGAGTACCCGAACTGCTGACAGACGACGACGATCTGTATCAACCAAGTAC 669

Qy 301 GACACCCGGAAGAAACCGCTTACTTTTGAACCTCGGAAAAAGACCGCAAGCCGATC 360

Db 670 GACACCTCGAAGAAACCGCTTACTTCTGAACTTTGAAAAAGACCGCAAGCCGATC 729

Qy 361 GTGCTGTGCTGCTTCATGCGCCCGGACCGCAATGTCGCGCAAGCATGCTCAACTG 420

Db 730 ATGCTGCTGCTTCATGCGCTTCGTGATACGCGCATGTCGCTGACGGGATGCTCAACTG 789

Qy 421 TACAAAGCTGCTGCTGCTGCGCAGCAACAGGACTGCGCGGCAAGGCGCTGCTGAC 480

Db 790 TACAAAGCTGCTGCTGCTGCGCAGCAACAGGACTGCGCGGCAAGGCTGCTGAC 849

Qy 481 ATGAAAGACGAGATCCAGTCGCGCGGAGAGTGAAGTGAATCAATCAAGACGGA 540

Db 850 ATGAAAGACGAGATCCAGTCGCGCGGAGAGTGAAGTGAATCAATCAAGACGGA 909

Qy 541 GCCTTCAAGAGCGCTGCGGCGCGCTGCGCATGCTGAGAGCAAGTCTGCTGCTTC 600

Db 910 GCATTTCAAGAGCGCTGCGGCGCGCTGCGCATGCTGAGAGCAAGTCTGCTGCTTC 969

Qy 601 CGCTGCGCGCAGAGCGCCACAGCGTCACTCGAGTTTCAATCAAGCAGATGACAGC 660

Db 970 CGCTGCGCTGCGCAGAGCGCCACAGCGTCACTCGAGTTTCAATCAAGCAGATGACAGC 1029

Qy 661 CTGCGCCAGGTGACATGCGCTTACAGCTATGAGCAAGCTCAACGACGAGCCCTCAAGGCC 720

Db 1030 CTGCGCAGCGTGAAGATGCGCTTACAGCTTACGAGCAAGCTGAGCGACACCGCTCAAGGCC 1089

Qy 721 CTGCGCAGACAGCGCGCCAAAGCGCTGATCCATCCCGGCAACCGGCAATGCTCGGTGCG 780

Db 1090 CTGCGCCAGAGCGCGTGGCAAGCATCATCATATCCGCGCAATGCTCGGTGCT 1149

Qy 781 TCGCGGCTGTGCGCAGCTTGCAGAGAGCTGCGCAAGACGCGTGCAGATCATTCGTTG 840

Db 1150 TCCAAAGTGTCTGCTGCTGCGCAGGAACTGCGCAAGACGCGTGCAGAAATCATTCGCTCT 1209

Qy 841 TCGC---GTCAACAGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897

Db 1210 TCCACGCTGAATGCGCGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1269

Qy 898 GACTGCTGCTGCGCCACGACTGAAACCGCAGAAAGCCCGCATCTGCGCATGTGTGCA 957

Db 1270 GACTGCTGCTGCGCCACGACTGAAACCGCAGAAAGCCCGCATCTGCGCATGTGTGCGC 1329

Qy 958 ATGACCAAGACCCGAGCAGCAGCAGAGAGCTGACGCGCATTTTCTGGGAATACTGA 1012

Db 1330 ATGACCAAGACCCGAGCAGCAGCAGAGAGCTGACGCGCATTTTCTGGGAATACTGA 1384

RESULT 6
PFL238710 2353 bp DNA linear BCT 30-SEP-1999
LOCUS Pseudomonas fluorescens endx gene.
DEFINITION AJ238710
ACCESSION AJ238710 GI:5002697
KEYWORDS endx gene; extracellular endonuclease.
SOURCE Pseudomonas fluorescens
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.

CR543861_34 3400001 3510000
CR543861_35 3500001 3598621
Continuation 721 of 36) of CR543861 from base 2000001 (CR543861 *Aginetobacter* sp. ADPI c

Query Match	30.9%	Score 313,	DB 1,	Length 110000,
Best Local Similarity	58.7%	Pred. No. 6.9e-29,		
Matches 582,	Conservative	0,	Mismatches 400,	Indels 9,
				Gaps 2,

QY 31 AACGTGGTGAATCCCTGGCCACCGGGCGGCACATCGCCGGGCGCTGGGCGCCAGGCGGGCCAAAC 90
 Db 71179 AACGTCGTTGTAAGTCGCAACAGGTGAAACAATGCGGGGTCCGGGCGCAAGCTCCACCAAC 71238
 QY 91 AGCGCCACTTACCAAGGCTGCCAAGGTTGCGCTGCACAAGACTGATTTGCGCGGTGCGAG 150
 Db 71239 AGTGCACCTTATCTGCGAGCAAAAGTACCAAGTCGATGCACTGATCTAAATGCAGTACTTCA 71298
 QY 151 CTGGCCGACCTTGGCCCATATGTGCGCGCGAGCAGGTGATGCAATGCGCTCCGAAGCATC 210
 Db 71299 ATCAAGATTTGGGCGCAATGTAAAGCGGCATTCAGAGCTTTACAAATCGCCTCTGAAGAGTATT 71358
 QY 211 ACCAAGCAGCAGCTGCTCAAGCTGGCAAGCAGCGCGGCGAGCTGGGCGGACGACGAATGAC 270
 Db 71359 ACTGACAAAGAAATTTATTTCTCTTGGCCGCAAGTTAAATGATCTTTGTTAAAAAACATCT 71418
 QY 271 GTCATGCGCATCGTCATCAACCATATGGCACCGACACCTTGGAAGAAACCGCTCACTTTTGG 330
 Db 71419 GTAAATAGGTGTTGTGATTAACAATGGTACAGATACGCTTGAAAGAAACAGCTTTCTTTT 71478
 QY 331 AACCTCGTGGAAAAAGACCGACAAAGCCGATCGTCTGTTGCTGTTCCATGCGCCCGGCAAC 390
 Db 71479 AATCTTGTGATTCTATACGAATAAACCAATTTCTGTTGTGTTCTATGCGCCCATCAACA 71538
 QY 391 GCCAATGTCCGCGCAGCGGATGCTCAACTGTATACACGCGCGTGGCGCGGCGCAGAACAAAG 450
 Db 71539 GCACCTTCAGAGATAGGCGCCCACTTAACCTTTATAGTGGCGGTTGCGCTTGTCTTCAGAT 71598
 QY 451 GACTCGCGCGGCAAGGCGGCTGCTGTGACATGAACGACGATTCAGATCCGCGCGTGCAC 510
 Db 71599 GAAACAAAAAATTAAGGCGGTATGGTATGTAATGATTCAACTTTTGACAGCAGTGAT 71658
 QY 511 GTGAGCAAGTGCATCAATCAATCAAGACCGAAGCTTCAAGAGCGCTGGGGCGCCGCTGGGC 570
 Db 71659 GTCAACAAAGGAATCAACATCATCAACATCAATTTGTATCAATGGGGGAAACTCGGT 71718
 QY 571 ATGTGTGTGAAGGCAAGTGCATCTGTTCCGCGCTGCGCGGCAAGCGCACACGATCAAC 630
 Db 71719 ACCTGCTGTGAAGGTAACTTCTGCTTTTGAAAGCTCAGTCAACGCGCATCAATGCG 71778
 QY 631 TCCGAGTTCCGACATCAAGCAGATCAGCAGC-----CTGCCCAAGTGGACATCGCTCAC 684
 Db 71779 TCTGAGTTAATATTGAAATAATCAAAAGGCGACCACTTCTCTGTTCAATGTGATAT 71838
 QY 685 AGCTATGGCAACGTCAACCGACCGCTTCAAGGCGCTTGGCACAGAACCGGCGCCAAAGCG 744
 Db 71839 GGTTCAGATCCCAAGTATCTGATGCGGTATCTTGCATATGCAAGGCGCAGGCGCTAAAGCG 71898
 QY 745 CTGATCCATGCGCGGCAACCGGCATGCGTCCGATGCTGCGGGGTGGTGGCGACCGCTGAG 804
 Db 71899 ATTATCAATGCTGTGTAACAGTAAATGCTCTGTAGTAACTAATATGTGCTCAACTTAAAG 71958
 QY 805 GAGCTGC---GCAAGAACCGCGGTGCAATCTGTTGCTGACGTCACACGAGCGCGTTTC 861
 Db 71959 AATATTACATGATGAAACAAGGGATTCAAATCAATCCGCTCATACGCGTATACGCAAGGTTTC 72018
 QY 862 GTGCTGCGTAAACGCTGAGCAGCCGACGACCAAGAACCAAGCACTGGGTCTGTTGGCCACGACTTG 921
 Db 72019 GTATTGTGTAATGCGGAACACCTGATGATCAATATGCTGGGTGTTGCACTAGATTTTA 72078
 QY 922 AACCCGACGAAGCGCCGATCCTGGCGATGTGTGGCAATGATCAAGAACCCAGAGCAGCAAG 981
 Db 72079 AATCCACAAAAGCAACGTCTGCTTGTGCTTTAGCACTGACAAAACCAATGATGTCCAA 72138
 QY 982 GAGCTGCAGCGCATTTTCTGGAAATACTGAT 1012

Db 72139 GAAATCCAGCGTATGTTCTGGCAATATTAAT 72169

RESULT 8
BX571965 16
WPCOMMENT

Sequence split into 41 fragments LOCUS BX571965 Accession BX571965

Fragment Name	Begin	End
BX571965_00	1	110000
BX571965_01	100001	210000
BX571965_02	200001	310000
BX571965_03	300001	410000
BX571965_04	400001	510000
BX571965_05	500001	610000
BX571965_06	600001	710000
BX571965_07	700001	810000
BX571965_08	800001	910000
BX571965_09	900001	1010000
BX571965_10	1000001	1110000
BX571965_11	1100001	1210000
BX571965_12	1200001	1310000
BX571965_13	1300001	1410000
BX571965_14	1400001	1510000
BX571965_15	1500001	1610000
BX571965_16	1600001	1710000
BX571965_17	1700001	1810000
BX571965_18	1800001	1910000
BX571965_19	1900001	2010000
BX571965_20	2000001	2110000
BX571965_21	2100001	2210000
BX571965_22	2200001	2310000
BX571965_23	2300001	2410000
BX571965_24	2400001	2510000
BX571965_25	2500001	2610000
BX571965_26	2600001	2710000
BX571965_27	2700001	2810000
BX571965_28	2800001	2910000
BX571965_29	2900001	3010000
BX571965_30	3000001	3110000
BX571965_31	3100001	3210000
BX571965_32	3200001	3310000
BX571965_33	3300001	3410000
BX571965_34	3400001	3510000
BX571965_35	3500001	3610000
BX571965_36	3600001	3710000
BX571965_37	3700001	3810000
BX571965_38	3800001	3910000
BX571965_39	3900001	4010000
BX571965_40	4000001	4074542

Continuation (17 of 41) of BX571965 from base 1600001 (BX571965 Burkholderia pseudomallea)

Query Match	Score	DB	Length
Best Local Similarity	27.9%;	283.2;	DB 1;
Matches	545;	Conservative	0;
		Mismatches	388;
		Indels	9;
		Gaps	2;

Query	DB	Score	DB	Length
24 GCTGGCAACGCGGATCTCTGCGCACCGGCGGCAACATCGCGGCGCTGCGCGCACGCGC	83	27.9%;	283.2;	DB 1;
83770 GCTGCGCGCATCGCGCTCTGCGCGCGGCGGCAAGATGCGGCGCGCGCGCGCGC	83829	27.9%;	283.2;	DB 1;
84 GCGCAACGCGCGCACCTACGCGCTGCGCAAGGTTGCGCTGCAACAGCTGATTCGCGCGCT	143	27.9%;	283.2;	DB 1;
83830 GCGCGACGCGCGCGCTACGCGCGCGCGCGCGCGCTGCGCTGCAAGCTGCTGCGCGCGT	83889	27.9%;	283.2;	DB 1;
144 GCGGAGCTGGCGCGCATCTGCGCGCATGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	203	27.9%;	283.2;	DB 1;
83890 GCGCGCGCTGCGCGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	83949	27.9%;	283.2;	DB 1;
204 AAGCATCACCAACGACGACCTGCTCAAGCTGCGCAAGCGAGCGAGCTGCGCGCGCGCGCAG	263	27.9%;	283.2;	DB 1;
83950 GGACTTGTGCGCGCGCGCTGTGACGACGCTTGGCGCGCGCGCGCGCGCGCGCGCGCG	84009	27.9%;	283.2;	DB 1;
264 CAATGACGTGATGCGCATGCTCATCAACCATGCGCACCGACCTCTGGAAGAAACCGCTTA	323	27.9%;	283.2;	DB 1;

QY 801 GACGAGCTGCGCAAGAACGCGGTGCAGATCATTCGTGTACGTCAACAGGGCGGTTT 860
 DB 68415 CCGCGATGCGCGCGGAGAGGCGGTGCGATGCTGCGCGCGTCCGCGCTGGGACA 68356
 QY 861 GGTCTGTGTAAACCGCGAGACCGCGGACGACAGAACGACTGGGTGTGGCCCAAGACT 920
 DB 68355 CGTATGTGGCAACGCGGCGGCGGACGACGATGCGCTCGGCTTCGTGAGCGCGGGTTCCT 68296
 QY 921 GAACCCGCAAGAGCGCGCATCTCTGCGCATGTGCGCATGAC 962
 DB 68295 CAATCCGTAAAGGCGCGGTGTGTGATGCTGCGCGTGC 68254

RESULT 10

LOCUS A27539 300 bp DNA linear PAT 20-SEP-1995
 DEFINITION Pseudomonas 7A glutaminase gene N-terminal.
 ACCSSION A27539
 VERSION A27539.1 GI:1247712

KEYWORDS
 SOURCE Pseudomonas sp.
 ORGANISM Pseudomonas sp.
 Bacteria; Proteobacteria.

REFERENCE 1 (bases 1 to 300)

AUTHORS Patent: DB 4140003-A 1 09-JUN-1993;

JOURNAL Location/Qualifiers

FEATURES
 source
 1..300
 /organism="Pseudomonas sp."
 /mol_type="unassigned DNA"
 /strain="ATCC 29598"
 /db_xref="taxon:306"
 <1..>300
 /codon_start=1
 /transl_table=1
 /product="7A-Glutaminase"
 /protein_id="CA01881.1"
 /db_xref="GI:1247712"
 /translation="KEYANOOKLANVYILATGTGTAGAGSANSATYQAAKGVDTL
 IAGVRELADLANVAGBQVMOJAKBSITNDLLKLASSVALADSNVDGIYTHGT"

CDS

Query Match 27.7%; Score 281; DB 6; Length 300;
 Best Local Similarity 96.0%; Pred. No. 2e-24;
 Matches 287; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 AAGGAATGAGAAACAGCAAGAGCTGCGCAAGCTGTGATCTGCGCAACCGCGGCACC 60
 DB 1 AAGGAATGAGAAACAGCAAGAGCTGCGCAAGCTGTGATCTGCGCAACCGCGGCACC 60
 QY 61 ATCCCGCGCGCTGCG 120
 DB 61 ATCCCGCGCGCTGCG 120
 QY 121 GTGCAAGAGTATGTCGAG 180
 DB 121 GTGCAAGAGTATGTCGAG 180
 QY 181 CAGGTGATGAGATGCTCCGAAAGCATCAACAGACGACTGTCTCAAGCTGGCAAGC 240
 DB 181 CAGGTGATGAGATGCTCCGAAAGCATCAACAGACGACTGTCTCAAGCTGGCAAGC 240
 QY 241 AGCGTGGCGAGCTGGCG 299
 DB 241 AGCGTGGCGAGCTGGCG 299

RESULT 11

LOCUS AR319786 1125 bp DNA linear PAT 17-AUG-2003
 DEFINITION Sequence 2336 from patent US 6562958.
 ACCESSION AR319786

VERSION AR319786.1 GI:33700889
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1125)
 AUTHORS Breton, G. and Bush, D.
 TITLE Nucleic acid and amino acid sequences relating to *Acinetobacter baumannii* for diagnostics and therapeutics
 JOURNAL Patent: US 6562958-A 2336 13-MAY-2003;
 FEATURES Location/Qualifiers
 source
 1..1125
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 26.9%; Score 273; DB 6; Length 1125;
 Best Local Similarity 56.2%; Pred. No. 1.3e-23;
 Matches 557; Conservative 0; Mismatches 425; Indels 9; Gaps 2;

QY 30 CAAGTGTGATTCCTGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 89
 DB 135 CAATGTTGTGTGCTGCTACTAGCGGTAACCATTTGCTGTGCGCGCAAGCTGACAAA 194
 QY 90 CAGCGCACCTACAGGCTGCGCAAGTGGCGTGCAGACAGCTGATTCGCGCGCGCGCGCA 149
 DB 195 TAGTGCACTTATACGCGCGCAAGTTCCAGTTGATGCTTATCAATGACAGTTCTCA 254
 QY 150 GCTGGCGGACCTGCGCAATGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 209
 DB 255 AATTCAGATTTGCGAGATGTATCTGGATTCAGACATTAACAAGTACTTCAAGAGAT 314
 QY 210 CACCAAGACGACCTGCTCAAGCTGCGCAAGCGGTGCGCGCGCGCGCGCGCGCGCATGA 269
 DB 315 TACTGACAGAGATATTATTAACATTTGCTGCAAGTAATTAATCTTTAAAGCCAAAC 374
 QY 270 CGTGATGAGCATGCTCATCAACCCATGCGACCGCACCCCTGGAAGAAACCGGCTACTTTT 329
 DB 375 TGTAAATGGCGTTGTGATTAACACGCGTACAGATCTTTAGAGAAACAGCATTTTCTT 434
 QY 330 GAACCTGTGAAAAAGACCGACAGCGCATGCTGTGCGGTTCATATGCGCGCGCGCGCAC 389
 DB 435 AATCTTGTGTTGATCATCTGATTAACCAATGTACTTGTGGCTCAATGCGCGCATCAAC 494
 QY 390 CGCCATGTCGCGCGCGCATGCTCAACCTGTACAGCGGTGCGCGCGCGCGCGCGCAACA 449
 DB 495 TGCTCTTCAGCAGATGTCACTTAACCTTTATATGCTGTTCATATGCGCGCGCGCTTCA 554
 QY 450 GGACTCGCGCGCGCAAGGCGGTGCTGTGACCATGACAGACAGATTCAGATCCGCGCGCTGA 509
 DB 555 TGATGCAAAAATTAAGGCGGTATGTGCTCATGACGACTTATTTTGTGCTGCTGA 614
 QY 510 CGTGAGCAAGTCATCAACATTAAGACCGAAGCCTTCAAGAGCGCTGCGCGCGCGCTGAG 569
 DB 615 TGTAACTTAAGGATTAACATTAACAAATGCTTTTGAAGCAATGCGGTGCTTTGGG 674
 QY 570 CATGTGTGTGAAGGCAAGTGATCTGCTTCCGCTCGCGCGCGCGCGCGCGCGCGCA 629
 DB 675 TACACTTGTGAAAGGCAACCATATTTGTTTGAACATCTGTAAAGTCAATCAATGTC 734
 QY 630 CTCGAGTTGACATCAAGGAGATCAGACG-----CTGCGCGCGGTGACATCGCTA 683
 DB 735 TTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 794
 QY 684 CAGCTATGCAACGTCAACGACCGGCTCAAGGCGCTGCGCAAGACGCGCGCAAGGC 743
 DB 795 TGCTTCACTATGCTTCCTGATGCTTAAGAGCATATGCAAAAGCTGGCGCTAAAGC 854
 QY 744 GCTGATCAATGCA 803
 DB 855 GATTATTAATGACAGGTACAGTATGCTTCTGAGCAAAATATATCTTCCATCAATTAACA 914
 QY 804 GAGCTGCG---GCAAGAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTTT 860

Db 915 AACCTTCATGACAAAACGGTATTCATTAATTCCTCATTCGCGCTTCACAAAGTTT 974
 Qy 861 CGGCTGCTGTAAGCGGAGGAGCCGACGACAAAGACGACGCTGCTGCGCCAGCACT 920
 Db 975 TGTATTCGATGATCAGAACCACTGACTTAATATGTTGGTGTAGACAGCTCATGATTT 1034
 Qy 921 GAACCCGACAGAGGCGCGCATCTCTGCGATGCTGCGCATATGACCAAGACCAAGACAA 980
 Db 1035 GAATTCCTCAAAAAGCGGTCTTCTTGCGGGGCTTGCGACTTACCAAAAACCAAGATGCGAA 1094
 Qy 981 GAAGCTGACAGCGCATTTTCTGGGAATCTGA 1011
 Db 1095 AGAATTCACAGCATGTTCTGGCAGTACTTA 1125

RESULT 12
 ECASN 1399 bp DNA linear BCT 31-MAR-1995
 LOCUS Erwinia chrysanthemi asn gene for L-asparagine aminohydrolyase (EC 3.5.1.1).
 ACCESSION X12746
 VERSION X12746.1 GI:40993
 KEYWORDS asn gene; hydrolase; L-asparagine aminohydrolyase.
 SOURCE Erwinia chrysanthemi (Pectobacterium chrysanthemi)
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Pectobacterium.
 REFERENCE 1 (bases 1 to 1399) Pullford, S. and Anderson, D.M.
 AUTHORS Filipula, D., Nagle, J.W., Pullford, S. and Anderson, D.M.
 TITLE Sequence of L-asparaginase gene from Erwinia chrysanthemi NCPPB 1125
 JOURNAL Nucleic Acids Res. 16 (21), 10385 (1988)
 MEDLINE 89057497
 PUBMED 3194219
 REFERENCE 2 (bases 1 to 1399)
 AUTHORS Filipula, D.
 TITLE Direct Submission
 JOURNAL Submitted (31-AUG-1988) Filipula D., Genex Corporation, 16020 Industrial Drive, Gaithersburg, MD 20877, USA

FEATURES
 source location/Qualifiers
 1..1399
 /organism="Erwinia chrysanthemi"
 /mol_type="genomic DNA"
 /strain="NCPB1125"
 /db_xref="taxon:556"
 109..115
 /note="inverted repeat A"
 123..129
 /note="inverted repeat A"
 209..214
 /note="poc. ribosome binding site"
 223..1269
 /codon_start=1
 /transl_table=1
 /product="put. L-asparagine aminohydrolyase precursor"
 /protein_id="CA31239.1"
 /db_xref="GI:40994"
 /db_xref="GOA:P06608"
 /db_xref="UniProt/Swiss-Prot:P06608"
 /translation="MERFSLFVLVLFPTASADKLPIVILATGTTASATG
 TQTTGKAGALGVDTLINAVEVKLANVKEQPSNNAENMGDVLKLSQRYNEL
 ARDDVDGVLTGTVESAYFLHLTVKSPKPVFVAARPAATASADPMNLAEV
 RVAGDKOSRGVWVINDRIGSARYITKTNASTLDPRANBEGYLVIIGNRIYGN
 RIKLHTTRGVDPRLGLSLPKVDLLGYDDPEPLYDAALQHGKGIIVVAGGAGSV
 SVRGIAKRLKLEKGVVVMSTRGNGIVPDEELPGLVSDSLNPAARILMLALTR
 TSDPKVIOEYFHY"

ORIGIN

Query Match 26.3%; Score 266.4; DB 1; Length 1399;
 Best Local Similarity 55.9%; Pred. No. 7.5e-23;
 Matches 557; Conservative 0; Mismatches 421; Indels 18; Gaps 2;

Qy 20 AGAAGCTGCGCAACGTGTGATCTTGGCCACCGGCGGACCAATCGCCGGCTGGGCCA 79
 Db 290 ATTAACCTGCCCAATATCGTTATCTTGGCCACCGGCGGTACTATGCGGCTGCGGCA 349
 Qy 80 GCGGCGCAACAGCGGCACCTTACAGAGTGGCCAGAGTGGCGTGCACAAGCTGATGGCG 139
 Db 350 CGGATACCCAAACACAGGCTTACAGAGCTGGCGCGCTTGGCTGGATACGCTTAATCAAG 409
 Qy 140 GCGTGCAGAGCTGCGCGACCTGGCCAAATGTCGCGCGAGAGGTATGACATGCGCT 199
 Db 410 CTGTGCTGAGTGAAGAACTGGCTTAATGTGAAGGGGAGGAGCTTCTCAACATGGCCA 469
 Qy 200 CCGAAGACATCACCAACAGACACTCTGACAGTGGGACAGACGCTGGCGGACTGGCCG 259
 Db 470 GCGAATAATGACCGGTGATGTGTCTCAAGCTGACCAAGCTGTGAATGAATCTGTGG 529
 Qy 260 ACAGCAATGACGTGATGGATGTCATGTCATGACCATGAGCCGACACCTTGAAGAAACG 319
 Db 530 CACGGATGATGTGATGTGTGTGTATCAACCGGTACCGACATGTTGGAAGAGTCGG 589
 Qy 320 CTTACTTTTGAACCTCTGGAAGAAAGACGACAGCCGATCTGTGTGCTTCAATGC 379
 Db 590 CTACTTTTCTTCACTGACGGTAAAGAGACAGCCGGTGTGTGTGTCAGCGATGC 649
 Qy 380 GCCCGGACCGGCATGTCCGCGACGCGCATCTCAACTGTACCAAGCCGTGGCGGTGG 439
 Db 650 GTCCGCAACAGGCGCATCAAGGCTGACGCGCCGACATGAACTGTGAAGCGGTGG 709
 Qy 440 CCAAGCAAGAGACTGCGCGGCAAGGCGGTGCTGATGACATGAAGCAAGATCCAGT 499
 Db 710 CGGATGATGACGTCTCGCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 769
 Qy 500 CCGGCGGTGACGTGACCAAGTGCATCAATCAAGCCGAAGCTTC--AAGAGCGCT 556
 Db 770 CTGCGGCTCATCAACCAAGACCAAGCCCTCTACGCTGATGATGTTCCGCGCAATGAAG 829
 Qy 557 GGGGCGCGCTGGGACATGTGTGTGAAGGCAAGTCTGATGCTTCCGCTCCGCGCAAGC 616
 Db 830 AAGGCTACCTGGGCGTTATCATGTGATGATGATGATGATGATGATGATGATGATGATGAT 889
 Qy 617 GCCACAGGTCACTCCAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 676
 Db 890 TGCATACCAACCGGT 949
 Qy 677 TCGCTTACAGCTTATGCAAGTGCACCGACCGCTTCAAGGCTTGGGACAGAGCGCG 726
 Db 950 TTTCTTATGCTATCAAGATGACCCGGAATATCTGTATGACGCGGCTATCACAGATGGCG 1009
 Qy 737 CCAAGCGCTGATCCATGCGCGGACCGGCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 726
 Db 1010 TAAAGGTATCGTCTATGCGGATGAGGCGCGGCTTCAAGTGTGTGTGTGTGTGTGTGTGTGT 1069
 Qy 797 CCGTGAAGAGTGTGCGGAAGAACGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 856
 Db 1070 GCATGCGCAAGGCGCTGGAAGAAAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1129
 Qy 857 GTTTCGTGCTGCTGAACCGCGGACGCGGACCGGACGACAAAGACGATGGGTGTGGCCACG 916
 Db 1130 GTATTGTGCGCGCGATGAAGAGCTGCGAG-----GTCGTGTGTGTGTGTGTGTGTGTGTGT 1174
 Qy 917 ACCTGAACCGCGAAGAGCGCGGATCTGTGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 976
 Db 1175 CTCTTAACCGCGGACATGCTGCGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1234
 Qy 977 GCAAGAGCTGACGCGCATTTCTGGGAATACTGAT 1012
 Db 1235 CGAAGTCATTCAAGATATTTCATATCTTATTGAT 1270

RESULT 13
 AY560098 1047 bp DNA linear BCT 22-MAR-2004
 LOCUS Erwinia chrysanthemi L-asparaginase precursor cds, complete cds.
 DEFINITION

ACCESSION AY560098
 VERSION AY560098.1 GI:45505306
 KEYWORDS
 SOURCE
 ORGANISM
 Brvlinia chrysanthemi (Pectobacterium chrysanthemi)
 Brvlinia chrysanthemi
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Pectobacterium.
 REFERENCE
 AUTHORS
 1 (bases 1 to 1047)
 Kozia, G.A. and Labrou, N.E.
 Cloning, expression and characterization of L-asparaginase from Brvlinia chrysanthemi 3937
 JOURNAL
 TITLE
 Unpublished
 2 (bases 1 to 1047)
 Kozia, G.A. and Labrou, N.E.
 REFERENCE
 AUTHORS
 Submitted (26-FEB-2004) Agricultural Biotechnology, Agricultural University of Athens, 75 Iera Odos, Athens 118 55, Greece
 JOURNAL
 TITLE
 Location/Qualifiers
 FEATURES
 source
 1..1047
 /organism="Brvlinia chrysanthemi"
 /mol_type="genomic DNA"
 /db_xref="taxon:556"
 CDS
 1..1047
 /function="catalyzes the hydrolysis of L-asparagine to L-aspartic acid and ammonia"
 /note="L-asparagine aminohydrolase"
 /codon_start=1
 /transl_table=1
 /product="L-asparaginase precursor"
 /protein_id="AA567028.1"
 /db_xref="GI:45505307"
 /translation="MERPKSLPIIAPRISTANADKLPIVILANGRTIAGSANG TQTGYKAGLGVDTLINAVPEYKLANVKEGPANASEMTDVTIKISQRNALL ARVDVGVITHTDTLESAYFLHLYTSKSPVPAAPAMPATVSLDGEMLIEAV RVADKQSRGVVAVLNDIRGSARYTKSNASTJDSFRANBEGYLVGIMHLYGN RLDKLTNRKSVDFVGRGLASLPKVDILVGYQDSELYLDAISHVKGIYVAGMGAGV SVRIAGRKAKODGIVVVRMSRTNGIVPDBALPGIVADSLNPAHRIILMALTR TSDPAVIGDYFHTY"
 819_peptide
 mat_peptide
 1..63
 64..348
 /product="L-asparaginase"
 ORIGIN
 Query Match 26.2%; Score 265.4; DB 1; Length 1047;
 Best Local Similarity 55.9%; Pred. No. 1,1e-22;
 Matches 556; Conservative 0; Mismatches 421; Indels 18; Gaps 2;
 20 AGAAGCTGGCCAAAGTGTGATCTCGGCGCACCGGCGGACCATGCGCGCGCTGGCGGCA 79
 68 ATAAGCTGCCAAATCTGCTATTCTGGCGACCGCGGCTGACTATTGCGCGGCTCGCGCGCA 127
 80 GCGGCGCCAAAGCGCCACTTACAGAGCTGCGCAAGTTGGCGTGCACAAAGCTGATTGCGG 139
 128 CGGGTACCCAGACTACCGGTTACAGAGCGCGCGCTGGCGTGGATACGCTGATCAACG 187
 140 GCGTGGCGGAGCTGGCGGCACTGGCCATGTGCGCGCGGAGACAGTATGATGCCCT 199
 188 CCGTGGCCGAGGTAAAAAACTGTCTAACGTAAAGGAGCAATTGCGCAACATGCGCA 247
 200 CCGAAGCATCATCAACGACGACCTGTCTCAAGCTGGCGCAAGCGGCGCGGCGGCGGCGG 259
 248 GCGAAAACATGACCGCGCATGTGTCTGAAACTGACGACGCGGTGATATCGCTGCTGG 307
 260 ACAGCAATGACGTGATGATGCTCATCAACCTTACGCAACGCAACCTGTGAGAAACCG 319
 308 CGCGTACGATGTGAGCGCGTGTGATTAACCAACGTAACGCAACGCTGAGAGATCGG 367
 320 CCTACTTTTGAACCTGTGAAAAAGACCGAACCGCATGTGCTGTGCTGCTTCATGC 379
 368 CCAATTTCTCTCACCTGACGGTGAAGGCAACGCGGTAGTGTGTTCGCGCGCGCATGC 427
 380 GCCCGGACACCGCATGTCCCGGACGCGCATGCTCAACTGTACAAAGCGCGTGGCGGTG 439

Db 428 GCCCGGCAAGCGCCATTAGAGCGCGACCGGCCGATGAACCTGTGAAAGCGGTGCGGTG 487
 Qy 440 CCAGCAACAGAGACTCGCGCGGCAAGGCGGTGTGTGACCATGAAGACAGATCCAGT 499
 Db 488 CGACGACAAACAGTCTCGCGCGCGCGCGGTGTGTGATGCTCAACGACCGTATCGGTT 547
 Qy 500 CCGGGCGGTGAGTGAAGATGATGATCAATCAACGACGAAAGCGTTCAA---GAGCGCT 556
 Db 548 CCGCGCGTTATATACCAAGACAAATGCTGAGCTGGAAGATTTCCGGCGGAATGAG 607
 Qy 557 GGGGCGCGTGGGATGTGTGTGAAAGCAAGTGTACTGTGTTCCGCTGCGCGCAAGC 616
 Db 608 AAGGTATCTGGGGTGTGTATCGGTACCATATCTATACAGAAATGCTTCAGAAAC 667
 Qy 617 GCGACAGGTCACTCCGAGTTTCAATCAACGATGACAGGCTTCCGAGGTGAC 676
 Db 668 TGCATACCAACGTTGCTGATTCATGTGCGCGCGCTGCTGCTGCGCAAGTCGATA 727
 Qy 677 TCGCTACAGCTATGGAAGATGACCAACGCGCTCAAGGCGCTGCGCAAGACGCG 736
 Db 728 TTTCTTATGGGTATCAAGATGATCTGGAATACCTGTACAGCGCGCATCAACGCG 787
 Qy 737 CCAAGCGCTGATCCATGCGCGCACCGCAATGCTGCTGCTGCGCGGTGTGCGAG 796
 Db 788 TGAAGGATGTTTATGCGCGTATGCGCGCGGTGTGATGCTGCGCGCATAGCG 847
 Qy 797 CCTGCAAGAGCTGCGCAAGACGCGCTGATGATCTTCTGTCACGTCAACAGGCG 856
 Db 848 GAATGCCCAAGCGCACGATTAAGCGGTGTGATGCGCTTCCTCCGCAACGCGCAACG 907
 Qy 857 GTTGTGCTGCGTAAAGCGCGACGCGCAACGCAAGACGATGCTGTGCGTGCACG 916
 Db 908 GCATGTGCGCGCGAAGAGCGCTGCGCGCC-----TGGTCCGATTT 952
 Qy 917 ACTGAACCGCGAAGAGCGCGCATCTGCGATGTGTGCAATGACCAAGACCGACGA 976
 Db 953 CGCTCAACCGCGCGCATGCGGATCTGTTGATGCTGCGCATGACCGACACGCGATC 1012
 Qy 977 GCAGAGCTGACAGCGCATTTCTGAGAAATCTGA 1011
 Db 1013 CGCGGTGATTCAGACTATTTCCACACTATTGA 1047
 RESULT 14
 BX640423 348251 bp DNA linear BCT 06-MAY-2004
 LOCUS
 DEFINITION
 1/14.
 BX640423 BX470249
 VERSION
 BX640423.1 GI:33564799
 KEYWORDS
 complete genome.
 SOURCE
 Bordetella parapertussis
 ORGANISM
 Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Alcaligenaceae; Bordetella.
 REFERENCE
 AUTHORS
 1
 Parkhill, J., Sebatia, M., Preston, A., Murphy, L.D., Thomson, N., Harris, D.E., Holden, M.T.G., Churcher, C.R., Bentley, S.D., Mungall, K.L., Cerdano-Tarraga, A.M., Temple, L., James, K., Harris, B., Quail, M.A., Achtman, M., Atkin, R., Baker, S., Basham, D., Bason, N., Cherevach, I., Chillingworth, T., Collins, M., Cronin, A., Davis, P., Doggett, J., Felwell, T., Goble, A., Hamlin, N., Haubert, H., Holt, R., Jorgensen, K., Leather, S., Mout, S., Norbertzak, H., O'Neill, S., Ormond, D., Price, C., Rabinowitsch, B., Rutter, S., Sanders, M., Saunders, D., Seeger, K., Sharp, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Unwin, L., Whitehead, S., Barrall, B.G. and Maskell, D.J.
 Comparative analysis of the genome sequences of Bordetella pertussis, Bordetella parapertussis and Bordetella bronchiseptica
 JOURNAL
 TITLE
 Unpublished
 2 (bases 1 to 348251)
 REFERENCE
 AUTHORS
 Sebatia, M.
 Direct Submission

JOURNAL Submitted (06-AUG-2003) Submitted on behalf of the Pathogen Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: ms5@sanger.ac.uk

FEATURES

Source

Location/Qualifiers

1..348251

/organism="Bordetella parapertussis"

/mol_type="genomic DNA"

/strain="12822"

/db_xref="taxon:519"

1..1920

/gene="gidA"

/locus_tag="BPP0001"

1..1920

/gene="gidA"

/locus_tag="BPP0001"

/note="Ortholog of Bordetella pertussis (BX470248) BPP0001 similar to *Bordetella pertussis* glucose inhibited division protein A GidA or B3741 SWALL:GIDB_ECOLI (SWALL:P17112) (629 aa) fasta scores: E(): 6.8e-164, 68.52% id in 629 aa"

/codon_start=1

/product="glucose inhibited division protein A"

/protein_id="CAE39742.1"

/db_xref="GI:33564800"

/db_xref="UniProt/TREMBL:Q7W211"

/translation="MDPREFDVIVGGHAGTEALAAAGAGQTLTLTHNTELLGQ MSCNPSTIGIGKHLVKEVDALGGAMATDEGIORINSSGPAVRATRYOADV GLYHAMPARLENONPMLFOOAVDDIMVOCDOVGATGILRPARATVLTGTFLN GLIHVGLQNTSGGADPPANSLQGRKEQLPQGRKTKTPPRIDORSINVTVEQ PGDIDVPVPSPLGKASMPHROQLPCTHTNTHNTHETIRGLDRSPMTSGVIGVGR YCPSIEDKIHIFADKASHQVLEBEGNTHETIYENGVTSLPDDVQELHSLGLEN AHIRPGVAYLEVDYPRALKSTLETVAISGLPACQINGTGYEEAAGLGAANA ALAOGKEPWPVRDEAVYGLVVDLVTGCTGTYEYRSTRAYRSLREDNDRLT EIGRRLGLVDDVPRDASRRKRDVAOVERLKTGWVRYLPAAHAEALIGKATERY SLSDLKRPVSYEALMQARTDEBELAGGVLEEDVLAQVETQYKATYKIQODE VQRLSHQQPIPADIDYDAVTSLSFEVRKRLKTHRPETIGQAAVSGVTPAISLL ILKRLHYGSRKQA"

19..1893

/gene="gidA"

/locus_tag="BPP0001"

/note="HMPfam hit to PF01134, Glucose inhibited division protein A, score 0"

820..864

/gene="gidA"

/locus_tag="BPP0001"

/note="P801280 Glucose inhibited division protein A family signature 1."

1917..2609

/gene="gidB"

/locus_tag="BPP0002"

1917..2609

/gene="gidB"

/locus_tag="BPP0002"

1917..2609

/gene="gidB"

/locus_tag="BPP0002"

/note="Ortholog of Bordetella pertussis (BX470248) BPP0002 similar to *Pseudomonas putida* glucose inhibited division protein B GidB SWALL:GIDB_PSEPU (SWALL:P25757) (216 aa) fasta scores: E(): 1.5e-24, 42.59% id in 216 aa, and to *Escherichia coli* glucose inhibited division protein B GidB or B3740 or B36462 SWALL:GIDB_ECOLI (SWALL:P17113) (207 aa) fasta scores: E(): 1.6e-12, 39% id in 200 aa"

/codon_start=1

/product="glucose inhibited division protein B"

/protein_id="CAE39743.1"

/db_xref="GI:33564801"

/db_xref="UniProt/TREMBL:Q7W210"

/translation="MSAVPDIPEGAPQRLACDARLPADAGQQKLTLYIEQMGRM NRYNLTALRDPQMLVOHLEDSLVAVGSGARVLPVGGSGGLPGV LAITRAMPVTCMDAVVEKTAFTVQMGALGIPYLQAAHRTIRLEBPACDDVYSRAF ASLDDPAKLAGRHVRBGGTLYVMKGVPPDEITQLQOHGHWYERIEPLVPALDAOR CLIMRRSSQGNL"

1917..2135

misc_feature

/gene="gidB"

/locus_tag="BPP0002"

/note="P800430 TonB-dependent receptor protein signature 1."

2004..2579

/gene="gidB"

/locus_tag="BPP0002"

/note="HMPfam hit to PF02527, Glucose inhibited division protein, score 1.7e-50"

2606..3403

/gene="para"

/locus_tag="BPP0003"

/note="synonym: soj"

2606..3403

/gene="para"

/locus_tag="BPP0003"

/note="Ortholog of Bordetella pertussis (BX470248) BPP0003 similar to *Caulobacter crescentus* chromosome partitioning protein Para or C63753 SWALL:PARA_CAUCU (SWALL:O05189) (267 aa) fasta scores: E(): 9.6e-46, 49.41% id in 267 aa, and to *Bacillus subtilis* sporulation initiation inhibitor protein Soj SWALL:SOJ_BACSU (SWALL:P37522) (253 aa) fasta scores: E(): 1.3e-43, 48.98% id in 247 aa"

/codon_start=1

/translation="Para family protein"

/product="Para family protein"

/protein_id="CAE39744.1"

/db_xref="GI:33564802"

/db_xref="UniProt/TREMBL:Q7W2H9"

/translation="MNIPSPKSARFPCANQKGVGKTTAINLAAGLTHKRVTL VDDPVGKNSGIDKSTLESYLVLGKAEITRTRSEGGVDVLPARELSGA EIDLVMQDREKQKKAIDKIEGVDVLDCEPTSLTLNLGAAHGVITPMOCEY FALGEGSDLVNTIKRHNINNEBLRYGLRFPDRMTLQOQVSAQLSHFQKVF TVPFRVRLAEAPVSGMPGVVDRASRAQAVIAFGAEMIERYEILD"

2879..3214

/gene="para"

/locus_tag="BPP0003"

/note="HMPfam hit to PF00991, Para family ATPase, score 1.5e-13"

3410..4216

/locus_tag="BPP0004"

3410..4216

/locus_tag="BPP0004"

/note="Ortholog of Bordetella pertussis (BX470248) BPP0004 similar to *Caulobacter crescentus* acetyltransferase, Gnat family CC2209 SWALL:G9A685 (EMBL:AB005892) (207 aa) fasta scores: E(): 1.5e-11, 34.63% id in 205 aa, and to *Rhizobium meliloti* putative acetyltransferase protein R02624 or SMC02449 SWALL:Q92MJ2 (EMBL:AL591791) (308 aa) fasta scores: E(): 5.5e-10, 31.42% id in 245 aa"

/codon_start=1

/translation="Putative acetyltransferase, Gnat family"

/product="Putative acetyltransferase, Gnat family"

/protein_id="CAE39745.1"

/db_xref="GI:33564803"

/db_xref="UniProt/TREMBL:Q7W2H8"

/translation="MPAVLAVLAHAPAPAGGLARLEBAALNATTVRQMLYEGMLV RMAPSRKACRQNVGLSTRPLDERLATCTRYAGHGVPLVRLTSGIDPFLDQL HERGFAPRPTCYMAPVVPFPAQASGLVETTFGPAQSGARLRGGDDATAEBA HRLQAAPLNTPLVADASGECTAAGLAAPDELIVCLPVLVTAPORRBOCGAAVLYX LLEQASHAGRQNTLOVEPONTARALYRFGVDVYTYWRMSATPV"

3929..4165

/locus_tag="BPP0004"

/note="HMPfam hit to PF00583, Acetyltransferase (GNAT) family, score 2.5e-14"

4260..5177

/gene="parB"

/locus_tag="BPP0005"

4260..5177

/gene="parB"

/locus_tag="BPP0005"

/note="Ortholog of Bordetella pertussis (BX470248) BPP0005 similar to *Pseudomonas putida* probable chromosome

misc_feature

/gene="gidB"

/locus_tag="BPP0002"

/note="P800430 TonB-dependent receptor protein signature 1."

2004..2579

/gene="gidB"

/locus_tag="BPP0002"

/note="HMPfam hit to PF02527, Glucose inhibited division protein, score 1.7e-50"

2606..3403

/gene="para"

/locus_tag="BPP0003"

/note="synonym: soj"

2606..3403

/gene="para"

/locus_tag="BPP0003"

/note="Ortholog of Bordetella pertussis (BX470248) BPP0003 similar to *Caulobacter crescentus* chromosome partitioning protein Para or C63753 SWALL:PARA_CAUCU (SWALL:O05189) (267 aa) fasta scores: E(): 9.6e-46, 49.41% id in 267 aa, and to *Bacillus subtilis* sporulation initiation inhibitor protein Soj SWALL:SOJ_BACSU (SWALL:P37522) (253 aa) fasta scores: E(): 1.3e-43, 48.98% id in 247 aa"

misc_feature

/gene="gidB"

/locus_tag="BPP0002"

/note="P800430 TonB-dependent receptor protein signature 1."

2004..2579

/gene="gidB"

/locus_tag="BPP0002"

/note="HMPfam hit to PF02527, Glucose inhibited division protein, score 1.7e-50"

/codon_start=1
/transl_table=11
/product="glucose inhibited division protein A"
/protein_id="CAE30503.1"
/db_xref="GI:33575040"
/db_xref="UniProt/TREMBL:Q7WRF1"
/translation="MDPFRFVDIVVGGHAGTEALAAARAGQTLTLTNITETGQ
MGNPSIGIGIKGHVKEVDALGAMAIADENGIOFRILNSKGPATRAQADRV
LYNNAIRLENPQNLMLFQAVDLMVQDDQVGAFTQIGLFRATVVLTAFTEN
GLHVLGIONVSGGRADPPANSLGORKELQLPQRLKTGPRIIDGRSINTVLEEO
PGDLDPVPSPTSGKASMPHROLPCWITHTNATHEITRGGLDRSPMSGIVGVEPR
YCSIEDKIHFRPADKASHOVFPBEPBELNTHETIPNVSSTSLPVDVYELIHSPGEN
AHLRPGVAVIETDFDPAKSLTETKALISGLFFAQINGTGYEPAAGLADANA
ALQAGKEPWPERRDAVIGVLDVIVTGVETPMTSRATYRLSLREDNADILYT
EIGRRLGLVDVWDVAFSRKRDVAQOEVRKSTWNPVPLPAAAEALIGKAIERYE
SLDLKRPVVSALMOARTDEGLAGVLEDAVQVETGVKAVGYARQODE
VOKHLSHEOQPIPADIDIDYAVTSLSEVNGKLTNHPRTIGQARVSGVTPAISLL
THKRLHYSRKKAA"
19. .1893
/gene="gida"
/locus_tag="BB0001"
/note="HMPfam hit to PF01134, DE glucose inhibited
division protein A, score 0"
820. .864
/gene="gida"
/locus_tag="BB0001"
/note="P801280 Glucose inhibited division protein A family
signature 1."
1810. .1875
/gene="gida"
/locus_tag="BB0001"
/note="Predicted helix-turn-helix motif with score
1129.000, SD 3.03 at aa 604-625, sequence
ERIGQARVSGVTPAISLLI"
1917. .2549
/gene="gida"
/locus_tag="BB0002"
/locus_tag="gidaB"
/note="ortholog of Bordetella pertussis (BX470248) BP0002"
/codon_start=1
/transl_table=11
/product="glucose inhibited division protein B"
/protein_id="CAE30504.1"
/db_xref="GI:33575041"
/db_xref="UniProt/TREMBL:Q7WRF0"
/translation="MSAVPDIPCGPQRLAQAQDALRLPADGQOQKLRITHEQWR
NRTYMLTALRDPQMLVQHLFDSLVAAPLERGLPGVLAIMKAMWDVTCVAVERT
AFVROMAGALGPNLQAATRIEQLPQAQDVVTSRAFSLQDPATLQGRHVEGGTL
VAMKGVPRDEIDALQOHGHWTVERIEPLVVPALDAQRCLIMWRBSQGN"
1917. .2135
/gene="gidaB"
/locus_tag="BB0002"
/locus_tag="gidaB"
/note="P500430 TonB-dependent receptor proteins signature
1."
2004. .2519
/gene="gidaB"
/locus_tag="BB0002"
/note="HMPfam hit to PF02527, DE glucose inhibited
division protein, score 4.7e-42"
2546. .3343
/gene="para"
/locus_tag="BB0003"
/locus_tag="para"
/note="synonym: boj"
2546. .3343
/gene="para"
/locus_tag="BB0003"
/locus_tag="para"
/note="ortholog of Bordetella pertussis (BX470248) BP0003"
/codon_start=1
/transl_table=11
/product="Para family protein"
/protein_id="CAE30505.1"

/db_xref="GI:33575042"
/db_xref="UniProt/TREMBL:Q7WRE9"
/translation="MKNIPPSKAVFVCIANQKGVGKTTAINLAAGLATHKORVLL
VLDPQGNATMGSIGDKSTLESNLVQVILGEAGIQTRVRSGGVDVLPARELSGA
EIDVOMERERQLKAIDKIGEDVFLVIDEPPTSLTLTLNGLAASHVILPMOCEY
PALBGLSLVNTIKRVHNNINLELVIGLRWPDPRMTLQOOVSAQESHSGDKVFT
TVPPNVLAEAPSTGMPGVYDRAKRAQAVIATGAEITERYEKL"
2819. .3154
/gene="para"
/locus_tag="BB0003"
/locus_tag="para"
/note="HMPfam hit to PF00991, DE Para family ATPase,
score 1.5e-33"
3350. .4156
/locus_tag="BB0004"
/locus_tag="paraB"
/locus_tag="BB0004"
/locus_tag="paraB"
/note="ortholog of Bordetella pertussis (BX470248) BP0004"
/codon_start=1
/transl_table=11
/product="Putative acetyltransferase, Gnat family"
/protein_id="CAE30506.1"
/db_xref="GI:33575043"
/db_xref="UniProt/TREMBL:Q7WRE8"
/translation="MPAVLAAPAPAPAGCEIARIBERAINATTREOMLYEGMLV
KMAPRARKACVNVLGSTRPDERLAYCTRYAGHGPVLIVRLTSLGPPDSLDQL
HERGFVAPDRTCVMAVPVFTAQAPSGVLFERTTPGRPAQAGRLRGHDNATASHA
AHLQAAPLENTPLVRDASGEICIAAGLAFDAELVGLFDIVTAPORRQCGCAAVLYK
ILEQASHAGARQAVYQVEPQNTAARALYRFQVRYTYVYRSMSEATPV"
3869. .4105
/locus_tag="BB0004"
/locus_tag="paraB"
/note="HMPfam hit to PF00583, DE Acetyltransferase (GNAT)
family, score 2.5e-14"
4200. .5117
/gene="paraB"
/locus_tag="BB0005"
/locus_tag="paraB"
/locus_tag="BB0005"
/locus_tag="paraB"
/note="ortholog of Bordetella pertussis (BX470248) BP0005"
/codon_start=1
/transl_table=11
/product="probable chromosome partitioning protein ParB"
/protein_id="CAE30507.1"
/db_xref="GI:33575044"
/db_xref="UniProt/TREMBL:Q7WRE7"
/translation="MATKKPKGLGRLDGLCADVPALDNICKAPAAPGEPATLVP
SMGRGKQOPRTRMDEGALNELAASIRTGIOPLIVRLGADAPQOYEIIIGERFR
AAOAGIKVNPVIVAEVDENAAWALIENTIGREDLNPLEAHGAVRLIDEGITHEO
MAAIGRSATSNIILRLINTLAAPVQTMILADVDVMGHARALADVAATQIQLANGYI
AARLVREAEKIVARAQKDVDAAPRKGGASRDVTRLEALSDHLGTRVALKVGARE
KQIIVDPFGWTHLSLIERQGLSGVDA"
4317. .4598
/gene="paraB"
/locus_tag="BB0005"
/locus_tag="paraB"
/note="HMPfam hit to PF02195, DE ParB-like nuclease
domain, score 5.9e-34"
4665. .4730
/gene="paraB"
/locus_tag="BB0005"
/locus_tag="paraB"
/note="Predicted helix-turn-helix motif with score
1337.000, SD 4.42 at aa 156-177, sequence
LTHQAAQAIGRSATSNTLNR"
5110. .6111
/gene="ansB"
/locus_tag="BB0006"
/locus_tag="ansB"
/locus_tag="BB0006"
/locus_tag="ansB"
/gene="ansB"
/locus_tag="BB0006"
/locus_tag="ansB"
/EC_number="3.5.1.1"
/note="ortholog of Bordetella pertussis (BX470248) BP0006"
/codon_start=1
/transl_table=11
/product="L-asparaginase II precursor"

```

/protein_id="CA30508.1"
/db_xref="UniProt/Tx:EMBL:O7WR86"
/translation="MPRARLPRIAVLATGCTIAGACQDDASGATRASVYGOULAAV
POLGLARIGAEQIANVGSQMSHAWRQLRLNALCDDPATDGIYTHGDTLEST
AYFLSLVLPDKPVLVGLAMRPATLADADGVNLVNLACHNAPAGRPVLYMND
IHLARHOKTAAAGSIIAASPANAGRAVMOGRPFVBARQADLLPDELPOQA
WPRVDIVHACVDLDGELIDPMAGRAOGIVLAGVDGNATPAALQALRASRGVAVVR
ASRTSGSRVGRNRIEIDDACGFIAGGLSLPQKANVLLTGLCQTRDRIARLQALPDSR"
mlsc_feature
5110..5172

```

```

Query Match 26.0%; Score 264; DB 1; Length 347356;
Best Local Similarity 56.5%; Pred. No. 3.6e-23;
Matches 555; Conservative 0; Mismatches 415; Indels 12; Gaps 3;

```

```

Qy 24 GCTGGCCAACTGATGATCTCTGGCCACCGGCGCAACCATCGCCGCGCTGGCCGCAAGCC 83
Db 5124 GCTGCCCGGATTCCTCTGGCGACCGGCGCACCATTCGCCGCGCGCAACCGCAAGCC 5183
Qy 84 GGCACACAGCGCACTACGAGCTGCGCAAGGTTGGCGTGAACAAGCTGATTGCGGGGT 143
Db 5184 GGCACAGCGGCT--TACGCGCGGCTGCTACGCGGTGCGGCAATTGCTGGCGCGGT 5240
Qy 144 GCGGAGCTGGCCGACCTGGCCAAATGCGCGGCGAGCAGGTGATGCAATCGCTCCGA 203
Db 5241 GCTCAACTGGCCGCGCTGGCGCGCATCGGCGCGAGAGATGGCCAAAGTGGCAGCCA 5300
Qy 204 AAGCATCAACCAACAACCTGCTCAAGCTGGCAAGCAGCTGGCCGAGCTGGCCGACAG 263
Db 5301 GAACATAGGCAATCCGCTGTCGCGGACGCTCGCGCGCGCGCGCTGTGCGACGA 5360
Qy 264 CAATGACGTGATGATGATGATCAACCATGAGCAGCAACCTGGAAGAAACGGCTA 323
Db 5361 TCGGCGACGACGCGCATGCTCATTCATCCATGACGATACGCTGAGAGAGACCGCTTA 5420
Qy 324 CTTTGAACCTCGTGAAAAAGACCAAGCCGATCGTGTGCTGCTTCATGCGCC 383
Db 5421 TTTCTCAGCTGTATGCTGTGATTAACCGGTGTGCTGTGGCGCGCATGCGCC 5480
Qy 384 CGGCAACCGCATGTCGCGCGCAAGCTGCTCAACCTGTACAAGCCGTCGCGCGCAG 443
Db 5481 GGCACCGCTGTGACCGCGCGCGCGCTCAATCTAACAACCGGTGTGCGCTTG 5540
Qy 444 CAACAAGACCTGCGCGGCAAGGCGGTGTGTGACCATGAAAGACGAGATCCAGTCCGG 503
Db 5541 CCAATCCGCGCGACGCGCGCGCGCGCGCTGTGTGATGAACAGAGATTCATCTGGC 5600
Qy 504 GCGTGAAGTGAAGTGAATCAATCAAGACCGAAGCCTTCAAGAGCGCTGGGCCC 563
Db 5601 CCGTCAATGTGAAGAAACGGCGCGCTGTGGCATGCGCGCTTGCTTGCCCAATGCAAG 5660
Qy 564 GCTGGGATGTGTGAGAGCAAGTGTACTGTTCGCTGCGCGCGCAAGCGCCAC 623
Db 5661 CCGGCGCGGTCAATGACGGGCGCGCGCGGTTCACGCGCGCGCGCGCGCGCGCG 5720
Qy 624 GGTCAACTCGAAGTTC-----GACATCAAGCAGATCAAGACCTGCGCCAGGTGACAT 677
Db 5721 CGCGGACGCGCTGCTCAGGCGCGAGCAATGCGCGCAGAGCGCTGCGCGCGGTGACAT 5780
Qy 678 CGCCTACAGCTATGCAACGTCAACGACAGCCTTAAGGCGCTGCGCAAGAACGCGCG 737
Db 5781 CGTGCAATGCTGTGTGATCTCGACGGCGAGCTCATGATTTTCAATGCGCGG---CGAGC 5837
Qy 738 CAAGGCGCTGATCATACCGCGCACCGGCAATGCTCGGTGCTGCGCGGTGTGCAAGC 797
Db 5838 CCAAGGCGATGCTGTGCGCGCGCGCGCGCGCGCGCAACCGCGCGCGCGCGCGCG 5897
Qy 798 CTTGCAAGAGCTGCGCAAGAACGCGCGTGAAGATCATTTGTTGTTCACTCAACAGGCGG 857
Db 5898 CTTGCAAGCGGCGCGCGCGCGCGCGGTGCGGTGTGCGCGCGCGCGCGCGCGCGCTG 5957
Qy 858 TTTGCTGCTGCTGAACCGCGCGAGCGCGCAACAGAACGACTGCGTGTGCGCGCAAG 917

```

```

Db 5958 CAGGTTGGGCGCAATATCGAATCGACGACGCGTGGCGCTTCATTCGCCCGCGCGA 6017
Qy 918 CTTGAACCCCGAAGAGCGCGCGCATCTCGCGATGTGGCAATGACCAAGCCAGACAG 977
Db 6018 CTTAGCGCCCGAAGAGCGCGCGCTGTGCTTACCTTGGCGCTGTGCGAAGCCGCGATAC 6077
Qy 978 CAAGAGCTGCAAGCATTTTC 999
Db 6078 CGCGCGCTGCAAGCGCTTGTTC 6099

```

```

Search completed: March 22, 2005, 11:48:42
Job time : 4641 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: March 22, 2005, 05:44:16 ; Search time 657 Seconds
(without alignments)
9136.404 Million cell updates/sec

Title: US-09-842-628-1
Perfect score: 1014
Sequence: 1 aaggaagtcgagaccagca.....tttctcgggaatcgtgataa 1014

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_GeneSeq_16Dec04:*
2: geneseqn1980s:*
3: geneseqn1990s:*
4: geneseqn2000s:*
5: geneseqn2001s:*
6: geneseqn2002s:*
7: geneseqn2003s:*
8: geneseqn2004s:*
9: geneseqn2005s:*
10: geneseqn2006s:*
11: geneseqn2007s:*
12: geneseqn2008s:*
13: geneseqn2009s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1014	100.0	1014	2	AAQ68438
2	742	73.2	1182	11	ABD07189
3	742	73.2	1135	11	ABD07273
4	562.2	55.4	2748	11	ABD07169
5	281	27.7	300	2	AAQ43365
6	273	26.9	1125	9	ADA31049
7	261.6	25.8	2837	1	AAAT0557
8	252.2	24.9	1044	3	AAK82834
9	240.8	23.7	1848	2	AAAT96346
10	240.4	23.7	1071	11	ACH94935
11	219	21.6	1133	2	AAK03474
12	219	21.6	1133	3	AAK62512
13	209.2	20.6	1174	8	ABZ80801
14	207.2	20.4	270	11	ABD07233
15	205.6	20.3	1303	12	ADQ07182
16	200.6	19.8	1330	12	ADQ07182
17	189.2	18.7	1128	12	ADJ36573
18	178	17.6	1137	12	ADN97129
19	167	16.5	1400	12	ADQ07180
20	158.2	15.6	110000	10	ACR67367_15

C	21	158.2	15.6	249878	10	ACR65381	ACR65381 Photorhab
	22	157.2	15.5	1044	10	ACR68782	ACR68782 Photorhab
	23	156.2	15.4	3223	12	ADN97128	ADN97128 ASPA01 as
	24	145.6	14.4	1186	2	AAK14328	AAK14328 H. pylori
	25	138	13.6	110000	2	AAK12063_07	Continuation (8 of
	26	138	13.6	110000	2	AAK12063_08	Continuation (9 of
	27	129	12.7	1041	10	ADP00934	AdP00934 Bacterial
	28	122.8	12.1	1470	12	ADQ07186	AdQ07186 Fusarium
	29	118.2	11.7	1260	12	ADQ07184	AdQ07184 Penicillium
	30	110	10.8	1236	12	ADQ07188	AdQ07188 Penicillium
	31	85.2	8.4	345	6	ABK6196	ABK6196 Helicobac
	32	83	8.2	110000	4	AAI99682_17	Continuation (18 o
	33	83	8.2	110000	4	AAI99683_17	Continuation (18 o
	34	82.6	8.1	990	12	ACH989418	ACH989418 Human gen
	35	80.2	7.9	3823	13	ADP07996	ADP07996 Pili leng
	36	78.8	7.8	2772	11	ABD07234	ABD07234 Pseudomon
	37	75.8	7.5	926	12	ACH87387	ACH87387 Human gen
	38	74.6	7.4	1191	8	ACA26934	ACA26934 Prokaryot
	39	72.2	7.1	960	10	ABX07792	ABX07792 S. pneumo
	40	72.2	7.1	960	12	ADN92005	ADN92005 S. pneumo
	41	72.2	7.1	966	13	ADR92156	ADR92156 Novel S.
	42	72.2	7.1	110000	10	ABSS6454_17	Continuation (18 o
	43	71.2	7.0	110000	6	ABA50521_07	Continuation (8 of
	44	70.6	7.0	2000	8	ADAT1938	ADAT1938 Rice gene
	45	70	6.9	2364	8	ABV74978	ABV74978 Synthetic

ALIGNMENTS

RESULT 1	AAQ68438	standard; DNA; 1014 BP.
ID	AAQ68438	
XX	AAQ68438;	
AC	25-MAR-2003 (revised)	
DT	12-JUN-1995 (first entry)	
XX	Pseudomonas glutaminase gene.	
DE	glutaminase; antiviral; virucide; anticancer; cancer therapy; HIV virus;	
XX	gene therapy; Escherichia coli; ds.	
OS	Pseudomonas sp.	
XX	WO9413817-A1.	
PN	23-JUN-1994.	
PD	04-DEC-1992; 92MO-US010421.	
PF	04-DEC-1992; 92MO-US010421.	
XX	04-DEC-1992; 92MO-US010421.	
PR	(MEME-) MB MEDICAL ENZYMS AG.	
PA	Roberts J, MacCallister TW, Sethuraman N, Freeman AG;	
XX	WPI; 1994-217891/26.	
DR	P-PDB; AAR59739.	
XX	Recombinant glutaminase derived from Pseudomonas 7A - expressed in E.	
PT	coli to increase yield and avoid Pseudomonas endotoxins for antiviral and	
PT	anticancer therapy.	
XX	Disclosure; Page 33; 60pp; English.	
PS	Chromosomal DNA from Pseudomonas sp. 7A (ATCC 29598) was used to	
CC	construct a genomic library in Escherichia coli LB32. Screening with	
CC	mixed oligonucleotide probes was used to isolate a glutaminase- encoding	
CC	clone. This was sequenced using the primers given in AAQ68439-47. The	
CC	gene can be used to manufacture recombinant glutaminase, free of	
CC	Pseudomonas exotoxin, for use in e.g. HIV and cancer therapy. The gene	

CC may also be used in gene therapy protocols. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 SQ Sequence 1014 BP; 226 A; 324 C; 318 G; 146 T; 0 U; 0 Other;

Query Match 100.0%; Score 1014; DB 2; Length 1014;
 Best Local Similarity 100.0%; Pred. No. 2,5e-166;
 Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAGGAAGTGAAGACGAGACGAGAGCTGGCCAAAGTGTATCTGTGACACCGGCGGACAC 60
DB 1 AAGGAAGTGAAGACGAGACGAGAGCTGGCCAAAGTGTATCTGTGACACCGGCGGACAC 60
QY 61 ATGCGCGGCGCTGGCCGACGCGGCGCAACGCGCACTACCAAGGCTGCAAGGTTGAC 120
DB 61 ATGCGCGGCGCTGGCCGACGCGGCGCAACGCGCACTACCAAGGCTGCAAGGTTGAC 120
QY 121 GTGCGACAGCTGATTTCCGCGCTGCGGAGCTGGCGGACCTGGCCATATGGCGCGGAG 180
DB 121 GTGCGACAGCTGATTTCCGCGCTGCGGAGCTGGCGGACCTGGCCATATGGCGCGGAG 180
QY 181 CAGGTGATGAGATGCGCTCCGGAAGCATCAACGAAGACCTGTCAAGCTGGCGAGC 240
DB 181 CAGGTGATGAGATGCGCTCCGGAAGCATCAACGAAGACCTGTCAAGCTGGCGAGC 240
QY 241 AAGCGTGGCCGAGCTGGCCGACGAGCAATGACGTGATGATGATCTCATCAACCCATGGCAC 300
DB 241 AAGCGTGGCCGAGCTGGCCGACGAGCAATGACGTGATGATGATCTCATCAACCCATGGCAC 300
QY 301 GACACCTCTGAAGAAACCGGCTACTTTTGAACCTGTGGAAGAAAGCCGACGATC 360
DB 301 GACACCTCTGAAGAAACCGGCTACTTTTGAACCTGTGGAAGAAAGCCGACGATC 360
QY 361 GTGCGTGTGGTTCATGAGGCGCGGACCGGACATGTCGCGGAGGAGATGCTCAACCTG 420
DB 361 GTGCGTGTGGTTCATGAGGCGCGGACCGGACATGTCGCGGAGGAGATGCTCAACCTG 420
QY 421 TACAAAGCCGCTGGCCGCTGGCCGACGAGCACTGCGCGGCAAGGCGCTGTGTGACC 480
DB 421 TACAAAGCCGCTGGCCGCTGGCCGACGAGCACTGCGCGGCAAGGCGCTGTGTGACC 480
QY 481 ATGAACGACGAGATTCAGTCCGGGCGTGAAGTGAAGATGATCAATCAAGACCGAA 540
DB 481 ATGAACGACGAGATTCAGTCCGGGCGTGAAGTGAAGATGATCAATCAAGACCGAA 540
QY 541 GCCTTCAAGAGCGCTGGGCGCGCTGGGCGATGCTGTGGAAGCAAGTGTGATGCTGTT 600
DB 541 GCCTTCAAGAGCGCTGGGCGCGCTGGGCGATGCTGTGGAAGCAAGTGTGATGCTGTT 600
QY 601 CGGCTGCGGCGCAAGGCGCAACGCTCACTCCGAGTTGCAATCAAGACGATCAGAC 660
DB 601 CGGCTGCGGCGCAAGGCGCAACGCTCACTCCGAGTTGCAATCAAGACGATCAGAC 660
QY 661 CTGCCCCAGGTGACATCGCTTACAGTATGAGCAAGTCAACGCGCTTCAAGAGGCG 720
DB 661 CTGCCCCAGGTGACATCGCTTACAGTATGAGCAAGTCAACGCGCTTCAAGAGGCG 720
QY 721 CTGCGACAGAAAGCGGCGCAAGGCGCTGATCAAGCGGCAACGCGCAATGCTGCGTGG 780
DB 721 CTGCGACAGAAAGCGGCGCAAGGCGCTGATCAAGCGGCAACGCGCAATGCTGCGTGG 780
QY 781 TCGCGGCTGGTGCACGCTTGCAGAGCTGCGCAAGAAAGCGGCTGCAATCTGTTG 840
DB 781 TCGCGGCTGGTGCACGCTTGCAGAGCTGCGCAAGAAAGCGGCTGCAATCTGTTG 840
QY 841 TCACTGCAACAGGCGGCTTCTGCTGCTGTAACGCGGAGCAAGCCGACGACGAAAGAC 900
DB 841 TCACTGCAACAGGCGGCTTCTGCTGCTGTAACGCGGAGCAAGCCGACGACGAAAGAC 900
QY 901 TGGGTGTGGCCCAACGATGAAACCGGCAAGAGGCGCATCTGTGCGATGTGGCAATG 960
DB 901 TGGGTGTGGCCCAACGATGAAACCGGCAAGAGGCGCATCTGTGCGATGTGGCAATG 960

```

QY 961 ACCAAGACCCAGAGACGACGAGAGCTGCAAGGCAATTTCTGGGAATACTGATTA 1014
 DB 961 ACCAAGACCCAGAGACGACGAGAGCTGCAAGGCAATTTCTGGGAATACTGATTA 1014

RESULT 2

ABD07189
 ID ABD07189 standard; DNA; 1182 BP.

XX ABD07189;

DT 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polynucleotide #5793.

DE Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;

KW antibacterial.

XX Pseudomonas aeruginosa.

OS US6551795-B1.

XX US6551795-B1.

PD 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfeld MJ, Nolling J, DeLoughery C, Bush D;

PI WPI; 2003-615309/58.

DR P-PsDB; ABO73618.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

XX useful as molecular targets for diagnostics, prophylaxis and treatment of

XX pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 5793; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the

XX polynucleotides encoding them. The sequences are useful in diagnosis and

XX therapy of pathological conditions, as molecular targets for diagnostics,

XX prophylaxis and treatment of pathological conditions resulting from a

XX bacterial infection, for evaluating a compound, such as a polypeptide,

XX for the ability to bind a P. aeruginosa nucleic acid, as components of

XX effective antibacterial targets, as targets for antibacterial drugs,

XX including anti-P. aeruginosa drugs, as templates for recombinant

XX production of P. aeruginosa-derived peptides or polypeptides, as target

XX components for diagnosis and/or treatment of P. aeruginosa-caused

XX infection, and in detection of P. aeruginosa sequences or other sequences

XX of Pseudomonas species using biochip technology. Sequences ABD01397-

XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:

XX The sequence data for this patent did not form part of the printed

XX specification but was obtained in electronic format from USPTO at

XX seqdata.uspto.gov/sequence.html

SQ Sequence 1182 BP; 247 A; 417 C; 361 G; 157 T; 0 U; 0 Other;

Query Match 73.2%; Score 742; DB 11; Length 1182;
 Best Local Similarity 83.9%; Pred. No. 2.6e-119;
 Matches 851; Conservative 0; Mismatches 160; Indels 3; Gaps 1;

```

QY 1 AAGGAAGTGAAGACGAGACGAGAGCTGGCCAAAGTGTATCTGTGACACCGGCGGACAC 60
DB 1 AAGGAAGTGAAGACGAGACGAGAGCTGGCCAAAGTGTATCTGTGACACCGGCGGACAC 60
QY 169 AAGGAAGTGAAGACGAGACGAGAGCTGGCCAAAGTGTATCTGTGACACCGGCGGACAC 228
DB 169 AAGGAAGTGAAGACGAGACGAGAGCTGGCCAAAGTGTATCTGTGACACCGGCGGACAC 228
QY 61 ATGCGCGGCGCTGGCCGACGCGGCGCAACGCGCACTACCAAGGCTGCAAGGTTGAC 120
DB 61 ATGCGCGGCGCTGGCCGACGCGGCGCAACGCGCACTACCAAGGCTGCAAGGTTGAC 120
QY 121 GTGCGACAGCTGATTTCCGCGCTGCGGAGCTGGCGGACCTGGCCATATGGCGCGGAG 180
DB 121 GTGCGACAGCTGATTTCCGCGCTGCGGAGCTGGCGGACCTGGCCATATGGCGCGGAG 180
QY 181 CAGGTGATGAGATGCGCTCCGGAAGCATCAACGAAGACCTGTCAAGCTGGCGAGC 240
DB 181 CAGGTGATGAGATGCGCTCCGGAAGCATCAACGAAGACCTGTCAAGCTGGCGAGC 240
QY 241 AAGCGTGGCCGAGCTGGCCGACGAGCAATGACGTGATGATGATCTCATCAACCCATGGCAC 300
DB 241 AAGCGTGGCCGAGCTGGCCGACGAGCAATGACGTGATGATGATCTCATCAACCCATGGCAC 300
QY 301 GACACCTCTGAAGAAACCGGCTACTTTTGAACCTGTGGAAGAAAGCCGACGATC 360
DB 301 GACACCTCTGAAGAAACCGGCTACTTTTGAACCTGTGGAAGAAAGCCGACGATC 360
QY 361 GTGCGTGTGGTTCATGAGGCGCGGACCGGACATGTCGCGGAGGAGATGCTCAACCTG 420
DB 361 GTGCGTGTGGTTCATGAGGCGCGGACCGGACATGTCGCGGAGGAGATGCTCAACCTG 420
QY 421 TACAAAGCCGCTGGCCGCTGGCCGACGAGCACTGCGCGGCAAGGCGCTGTGTGACC 480
DB 421 TACAAAGCCGCTGGCCGCTGGCCGACGAGCACTGCGCGGCAAGGCGCTGTGTGACC 480
QY 481 ATGAACGACGAGATTCAGTCCGGGCGTGAAGTGAAGATGATCAATCAAGACCGAA 540
DB 481 ATGAACGACGAGATTCAGTCCGGGCGTGAAGTGAAGATGATCAATCAAGACCGAA 540
QY 541 GCCTTCAAGAGCGCTGGGCGCGCTGGGCGATGCTGTGGAAGCAAGTGTGATGCTGTT 600
DB 541 GCCTTCAAGAGCGCTGGGCGCGCTGGGCGATGCTGTGGAAGCAAGTGTGATGCTGTT 600
QY 601 CGGCTGCGGCGCAAGGCGCAACGCTCACTCCGAGTTGCAATCAAGACGATCAGAC 660
DB 601 CGGCTGCGGCGCAAGGCGCAACGCTCACTCCGAGTTGCAATCAAGACGATCAGAC 660
QY 661 CTGCCCCAGGTGACATCGCTTACAGTATGAGCAAGTCAACGCGCTTCAAGAGGCG 720
DB 661 CTGCCCCAGGTGACATCGCTTACAGTATGAGCAAGTCAACGCGCTTCAAGAGGCG 720
QY 721 CTGCGACAGAAAGCGGCGCAAGGCGCTGATCAAGCGGCAACGCGCAATGCTGCGTGG 780
DB 721 CTGCGACAGAAAGCGGCGCAAGGCGCTGATCAAGCGGCAACGCGCAATGCTGCGTGG 780
QY 781 TCGCGGCTGGTGCACGCTTGCAGAGCTGCGCAAGAAAGCGGCTGCAATCTGTTG 840
DB 781 TCGCGGCTGGTGCACGCTTGCAGAGCTGCGCAAGAAAGCGGCTGCAATCTGTTG 840
QY 841 TCACTGCAACAGGCGGCTTCTGCTGCTGTAACGCGGAGCAAGCCGACGACGAAAGAC 900
DB 841 TCACTGCAACAGGCGGCTTCTGCTGCTGTAACGCGGAGCAAGCCGACGACGAAAGAC 900
QY 901 TGGGTGTGGCCCAACGATGAAACCGGCAAGAGGCGCATCTGTGCGATGTGGCAATG 960
DB 901 TGGGTGTGGCCCAACGATGAAACCGGCAAGAGGCGCATCTGTGCGATGTGGCAATG 960

```

```

OY 121 GTGCAAGAGTATGTCGCGGCGGAGATGCGGACCTGCGCAATATGCGCGGCGAG 180
DB 289 GTGCAAGAGTATGTCGCGGCGGAGATGCGGACCTGCGCAATATGCGCGGCGAG 348
OY 181 CAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
DB 349 CAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 408
OY 241 AGCGTGGCGGAGTGGCGGAGCAATGAGTGGATGATGATGATGATGATGATGATG 300
DB 409 ACCGTGGCGGAGTGGCGGAGCAATGAGTGGATGATGATGATGATGATGATGATG 468
OY 301 GACACCTGGAAGAAACCGCTACTTTTGAACCTTGTGAAAGAAAGCAAGCCGATC 360
DB 469 GACACCTGGAAGAAACCGCTACTTTTGAACCTTGTGAAAGAAAGCAAGCCGATC 528
OY 361 GTGTGTGTGATTCATGTCGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAG 420
DB 529 GTGTGTGTGATTCATGTCGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAG 588
OY 421 TACAAAGCGGTGGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCAATGTC 480
DB 589 TACAAAGCGGTGGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCAATGTC 648
OY 481 ATGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 649 ATGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
OY 541 GCGTTCAAGAGCGGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCAATGTC 600
DB 709 GCGTTCAAGAGCGGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCAATGTC 768
OY 601 CCGCTGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCA 660
DB 769 CCGCTGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCA 828
OY 661 CTGCGCGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 829 CTGCGCGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888
OY 721 CTGCGCGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 889 CTGCGCGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
OY 781 TCGCGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 949 TCGCGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008
OY 841 TCAC---GTCAACAGGCGGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 897
DB 1009 TCGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
OY 898 GACTGAGTGTGTCGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCAATGTCGCG 957
DB 1069 GACTGAGTGTGTCGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCAATGTCGCG 1128
OY 958 ATGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCA 1011
DB 1129 ATGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCAAGACCA 1182

```

RESULT 3

ABD07273/c
ABD07273 standard, DNA, 1335 BP.

AC ABD07273;
XX
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polynucleotide #5877.
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.

```

XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
DR WPI; 2003-615309/58.
DR P-PSDB; ABO73702.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 5877; 455bp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 1335 BP; 189 A; 404 C; 485 G; 257 T; 0 U; 0 Other;

```

Query Match 73.2%; Score 742; DB 11; Length 1335;
Best Local Similarity 83.9%; Pred. No. 2.6e-119;
Matches 851; Conservative 0; Mismatches 160; Indels 3; Gaps 1;

```

OY 1 AAGGAAGTGGAGAACGACGAGAGTGGCAAGTGGATCTGAGCCACGCGCGGAC 60
DB 1239 AAGGAAGTGGAGAACGACGAGAGTGGCAAGTGGATCTGAGCCACGCGCGGAC 1180
OY 61 ATGCGCGGCGTGGCGGAGCGGAGCAAGCGGACCTTACAGAGTGGCG 120
DB 1179 ATGCGCGGCGTGGCGGAGCGGAGCAAGCGGACCTTACAGAGTGGCG 1120
OY 121 GTGCAAGAGTATGTCGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCA 180
DB 1119 GTGCAAGAGTATGTCGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCA 1060
OY 181 CAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
DB 1059 CAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1000
OY 241 AGCGTGGCGGAGTGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCAATGTCGCG 300
DB 999 AGCGTGGCGGAGTGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCAATGTCGCG 940
OY 301 GACACCTGGAAGAAACCGCTACTTTTGAACCTTGTGAAAGAAAGCAAGCCGATC 360
DB 939 GACACCTGGAAGAAACCGCTACTTTTGAACCTTGTGAAAGAAAGCAAGCCGATC 880
OY 361 GTGTGTGTGATTCATGTCGCGGCGGAGCAATGTCGCGGAGCAATGTCGCGGAGCA 420

```

[illegible]

RESULT 4	
ABD07169	
ID	ABD07169 standard; DNA; 2748 BP.
XX	
AC	ABD07169;
XX	
DT	29-JUL-2004 (first entry)
XX	
DB	Pseudomonas aeruginosa polymnucleotide #5773.
XX	
KK	Bacterial infection; gene; ds; Pseudomonas aeruginosa infection,
KW	antibacterial.
XX	
OS	Pseudomonas aeruginosa.
XX	
PN	US6551795-B1.
XX	
PD	22-APR-2003.
XX	
PF	18-FEB-1999; 99US-00252891.
XX	
PR	18-FEB-1998; 98US-0074788P.
XX	
PR	27-JUL-1998; 98US-0094190P.
XX	
PA	(GENO-) GENOME THERAPEUTICS CORP.
XX	
PI	Rubenfield M7, Nolling J, Deloughery C, Bush D;
XX	
DR	WPI; 2003-615309/58.

DR P-PSDB; ABO73598 .
XX
PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
PT useful as molecular targets for diagnosis, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 5773; 455bp; English.
XX
XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnosis,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biotech technology. Sequences ABB01397-
CC ABB17967 represent *P. aeruginosa* polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification and was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 2748 BP; 504 A; 1016 C; 870 G; 358 T; 0 U; 0 Other;

Query Match	55.4%	Score 562.2	DB 1	Length 2748
Best Local Similarity	85.2%	Pred. No. 3e-88		
Matches 640	Conservative 0	Mismatches 108	Indels 3	Gaps 1
QY	264	CAATGACGTGATGAGCATGTCATCAACCATGAGCAACCGACACCTCTGAGAGAAACCGCTTA	3233	
Db	1	CGACGACGTGCGACGGGATGTCATCAACCGATACCGATACCGACCTCTGAGAGAACCGCTTA	60	
QY	324	CTTTTGAACCTTCGTGGAGAAAGACCGACAAACCGATGTCGTGGTCCGTTCCATGCGCCC	3833	
Db	61	CTTCCCTGACCCCTGTCGTGAGCACACCGAAGACCTTATCGTGTGTCGTGCTCATGCGCCC	120	
QY	384	CGGCAACGGCATGTCGCGCCGACGGGCAATCTGAACCTGTACAAACGCGCGTGGCCGTGGCAG	4433	
Db	121	GGGCAACCGGATGTCGCGCCGACGGGCAATCTGAACCTGTACAAACGCGCGTGGCGTGGCCG	180	
QY	444	CAACAAAGACTCGCGCGGCAAGGCGGTGTCGTGACCAATGAACGACGAGATCCAGTCCGG	5033	
Db	181	CGACAAATGTCGCGACCGGGCAAGGCGGTGTCGTATCCATGAACGACGAGATCCCTCCGG	240	
QY	504	GGCTGACGTGAGCAAGTCGATCAATCAATCAAGCCGAAGCCTTCAAGAGCCCTGAGGCCC	5633	
Db	241	CCCGACCGGACCAAGATGCTCAATCAATCAAGCCGAAGCCTTCAAGAGCCCTGAGGCCC	3000	
QY	564	GCTGGGCATGTCGTGTCGAAAGCAAGTCGATCTGGTCCGCTGCGGACCAAGCGSCACAC	6233	
Db	301	GCTGGGCATGTCGTGTCGAGGSCAAGAGCTACTGGTCCGCGCACCGGTGAAGCGGACAC	360	
QY	624	GATCACTCCGAGTTGACATCAATCAAGCAGATCAGACGCTGCGCCGAGGTGACATGCTTA	6833	
Db	361	GATGAACTCCGAGTTGACATCAATCAAGCAGATCTCCGCGCTGAGCTCCGGTGAATACTGCTTA	420	
QY	684	CAGCTATGGCAACGTTCACCGACACGGGCTTAAGAGCCCTTGGACACAGAAACGGCGTCMAAGGC	7433	
Db	421	CAGCTATGGCAACGTTCAGGACACCGGCTTAAGAGCCCTTGGACACAGAAACGGCGTCMAAGGC	480	
QY	744	GCTGATCAATGCGGACACCGGCAATGCGTCCGATGTCGTGCGGAGTGGTGCAGGCGCTGCA	8033	
Db	481	GATCATCAATGCGGACACCGGCAATGCGTCCGATGTCGTGCGGAGTGGTGCAGGCGCTGCA	540	
QY	804	GGAGTCGCGCAAGAACGGGCTGTCAGATCATTTCTTCTGAC---GTCAACAGGCGGATTT	860	
Db	541	GGAGTCGCGCAAGAACGGGCTGTCAGATCATTTCTTCTGAC---GTCAACAGGCGGATTT	600	
QY	861	CGTCTGCGTAACGCCGACAGCCCGACGACAGACATCTGGGTCGTGGCCCAACACTT	9220	

Db 601 CGTCTGCGCAACGCCGAGACGCCGAGACAGACGATGTCGTGCCACGACCT 660
 QY 921 GAACCCGAGAAAGCCCGCATCTGTCGATGTGTCGATGACCAAGACCCAGACGAA 980
 Db 661 GAACCCGAGAAAGCCCGCATCTGTCGATGTGTCGATGACCAAGACCCAGACGAA 720
 QY 981 GGAGCTGCAAGCGCATTTTCTGAGATATCTGA 1011
 Db 721 GGAGCTGCAAGCGCATTTTCTGAGATATCTGA 751

RESULT 5
 AAQ43365
 ID AAQ43365 standard; DNA; 300 BP.
 XX
 AC AAQ43365;
 XX
 XX 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 13-OCT-1993 (first entry)
 XX
 DE Pseudomonas 7A glutaminase N-terminus coding region.
 XX
 KW Neoplasm; glutamine; anticancer; antiviral; therapy;
 KW Pseudomonas 7A-Glutaminase-Asparaginase; PGA; ss.
 XX
 OS Pseudomonas .sp. 7A.
 XX
 XX
 FH Key Location/Qualifiers
 FT mat_peptide 1..300
 FT /tag= a
 FT /note= "corresponds to first 100 amino acids at mature N-terminus of enzyme"
 FT 200..201
 FT /*tag= b
 FT /note= "not identified"

DE4140003-A1.
 XX
 PN 09-JUN-1993.
 PD
 XX
 XX 04-DEC-1991; 91DE-04140003.
 PP
 XX
 XX 04-DEC-1991; 91DE-04140003.
 PR
 XX
 PA (ROBE/) ROBERTS J.
 XX
 PI Robert J.
 XX
 DR WPI; 1993-189322/24.
 DR P-PSDB; AAR37660.
 XX
 PT human tissues or fluids for treating cancer and HIV.
 PT
 PT human tissues or fluids for treating cancer and HIV.
 PS
 PS Claim 5, Fig 1, 16pp; German.
 XX
 XX A gene library was prepared from Sau3A-digested chromosomal DNA of
 CC Pseudomonas 7A (ATCC 29598). Three oligonucleotide probes were
 CC synthesized based on amino acid sequences of peptide fragments from the N-
 CC terminus, middle and C-terminus of the mature glutaminase (see AAQ43362-
 CC Q43364, respectively). A single clone hybridized to all three probes.
 CC Further analysis resulted in isolation of a 1.1kb SalI fragment which
 CC encoded a 23 amino acid signal peptide and some of the N-terminal amino
 CC acids of the mature protein. Sequence AAQ43365 starts from the mature N-
 CC terminus of the glutaminase. The enzyme will be potentially useful in
 CC anti-tumor therapy in neoplastic tissues where glutamine levels are
 CC often increased. (Updated on 25-MAR-2003 to correct PN field.) (Updated
 CC on 24-OCT-2003 to standardise OS field)
 CC
 XX Sequence 300 BP; 66 A; 95 C; 97 G; 40 T; 0 U; 2 Other;
 SQ
 Query Match 27.7%; Score 281; DB 2; Length 300;

Beet Local Similarity 96.0%; Pred. No. 1.3e-39;
 Matches 287; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 1 AAGAAAGTGAAGAACACAGACAGAAAGTGGCAACGTGTGATCTCTGCGCAACCGCGGCAAC 60
 Db 1 AAGAAAGTGAAGAACACAGACAGAAAGTGGCAACGTGTGATCTCTGCGCAACCGCGGCAAC 60
 QY 61 ATCCGCGGCGCTGGGCGCGGCGGCGCAAGCGGCACTTACAGAGCTGCCAAGTTGAC 120
 Db 61 ATCCGCGGCGCTGGGCGCGGCGGCGCAAGCGGCACTTACAGAGCTGCCAAGTTGAC 120
 QY 121 GTCCAGAAAGTGAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 Db 121 GTCCAGAAAGTGAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 QY 181 CAGGTGATGCAATGCTCTCCGAAAGCATCACCAACGACATGCTCTCAAGTGGCAAGC 240
 Db 181 CAGGTGATGCAATGCTCTCCGAAAGCATCACCAACGACATGCTCTCAAGTGGCAAGC 240
 QY 241 AGCGTGGCGGAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 299
 Db 241 AGCGTGGCGGAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 299

RESULT 6
 ADA31049
 ID ADA31049 standard; DNA; 1125 BP.
 XX
 AC ADA31049;
 XX
 XX 20-NOV-2003 (first entry)
 DT
 DT
 DE DNA encoding Acinetobacter baumannii protein #2336.
 XX
 KW ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
 KW vaccine; plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 OS
 XX
 XX US6562958-B1.
 PN
 XX
 PD 13-MAY-2003.
 PP
 XX
 XX 04-JUN-1999; 99US-00328352.
 PR
 XX
 XX 09-JUN-1998; 98US-0086701P.
 PR
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI; 2003-576092/54.
 DR P-PSDB; ADA35175.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 PS
 PS Example; SEQ ID NO 2336; 328bp; English.
 CC
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species, in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents DNA encoding an A. baumannii
 CC protein.
 CC
 XX Sequence 1125 BP; 333 A; 224 C; 223 G; 345 T; 0 U; 0 Other;
 SQ

Query Match 26.8%; Score 273; DB 9; Length 1125;
 Best Local Similarity 56.2%; Pred. No. 3e-38;
 Matches 557; Conservative 0; Mismatches 425; Indels 9; Gaps 2;

```

QY 30 CAAGTGTGATTCCTGGCCACCGCGGACACATCGCCGCGCGCCGCGCCGCGCCAA 89
DB 135 CAATGTGTGTGTGCTGCTACTGGCGGTACATGTCTGTCTGTCCGCAACCTCAGCAAA 194
QY 90 CAGCGCCACCTACACGAGCTGCGCAAGGTTGGCGTCGACACAGCTGATGTCGCGCGTCCGA 149
DB 195 TAGTGCAACTTATACGCGCGCAAAAGTTCCAGTTGATCTTATTCATGACAGTTTCTCA 254
QY 150 GCTGCGCAGCTGGCCAAATGTGCGCGCGACGAGGTGATGCAATCGCTCCGAAAGCAT 209
DB 255 AATTCAAGATTGGCGAATGTATCTGTATCCAGCATTAACAGAGAGCTTCTGAAAGTAT 314
QY 210 CACCAAGACGACCTGCTCAAGCTGGCAAGAGGTGGCCGAGCTGGCGCGACAGCAATGA 269
DB 315 TACTGACAAAGAAATTTATTAATAATTTGCTGCAGATAATTAACCTGTTAAAGCCAAAC 374
QY 270 CGTCGATGGCATTCGTCATACCCATGCGACCGACACCTCGAAGAAACCGCTACTTTT 329
DB 375 TGTAAATGCGCTTGATTAACAACGCTACAGATTAATTAAAGAAACAGATTTTCTT 434
QY 330 GAACCTCGTGAAGAAACCGACGACGATCGTGTGCTGCTTCATGCGCCCGCGCAC 389
DB 435 AATCTGTGTCTCATCTGATTAACCAATCGTACTTGTGGCTCAATGCGCCCATCAAC 494
QY 390 CGGCATGCGCCGCGACGCGCATGCTCAACCTGTACAAAGCCGTGGCCGCGCGACGACAA 449
DB 495 TGCCTTTTCACAGATGTGTCACTTAACCTTTATAGTCTGTGATGTGCGCGCTTCTGA 554
QY 450 GGACTCGCGCGGCAAGGCGGTGCTGTGACCATGACGACGACAGATTCAGTCCGCGCTGA 509
DB 555 TGATGCAAAAATTAAGCGCTTATGCTTCTCATGACGACTTATTTTGTGCTCGTGA 614
QY 510 CGTGAAGCATGTCATCAACATCAAGCCGTTCAAGAGCGCTGCGCGCTGCGTGG 569
DB 615 TGTAACTTAAGGCAATTAATTCATTAACAATGCTTTTGAAGCCAAATGGCGTGTGG 674
QY 570 CATGTGTGTGAAGGCAAGGTGTAAGTGTGCGCTGCGCGCGCAAGCGCCACAGGTGAA 629
DB 675 TACCTGTGTGAAGGCAACATATGTTAGAACATCTGTTAAAGTCAACCAATGC 734
QY 630 CTCCGAGTTCCACATCAAGCATCAGCAGC-----CTGCCCGAGGTGACATCGCCTA 683
DB 735 TTCAGAATTTATATTTGAATTAATTAAGGTGATGCTTTCGACAGTACAAATCGTTTA 794
QY 684 CAGCTATGGAAGTCAACGACGACGCGCTTACAGAGCCCTGGGACAGAAAGCGCCAGGC 743
DB 795 TGGTTCGACTCTATGCTTCTGATGCTTATGAAGCATATGCAAAAGCTGCGCTAAAGC 854
QY 744 GGTATTCATGCGCGGACGCGCAATGCGTGTGCTGCGCGGAGTGGCGCGAGCCCTGCA 803
DB 855 GATTATTCATGAGGTACAGGTATGTTCTGTAGCGAAATATATCGTTCAACATTAACA 914
QY 804 GGAAGTGC---GCAAGAACGCGGTGACAGATTCGTTCTGACGTCACAGCGCGGTTT 860
DB 915 AAACCTTCATGCAAAAACGATTAATCAATTAATCCGCTCATCGCGCTTCACAAAGTTT 974
QY 861 CGTGTGTGTAACGCGCGACGCGCGACGACGACAGAACGATGCGGTGCTGCGCCAGACT 920
DB 975 TGTATTCATGATGACAGAACCTGACTTAATATATGTTGGGTAGGAGCTCATGATTT 1034
QY 921 GAACCCGAGAAAGCGCGCATCTGCGCGATGTGGCAATGAACAGACCGCAGACAGAA 980
DB 1035 GAATCTCTAAAGAGCGCTTCTTGGCGCGTGGCACTTACAAAGCAACGATGCGAA 1094
QY 981 GGAGCTGACGAGCTTTTCTGGGAATCTGA 1011
DB 1095 AGAATTCACGCAATGTTCTGGCACTACTAA 1125

```

RESULT 7
 AAN70557
 ID AAN70557 standard; DNA; 2837 BP.

```

XX AC AAN70557;
XX AC 25-MAR-2003 (revised)
XX DT 07-AUG-1991 (first entry)
XX DE Sequence encoding a protein with L-asparaginase (LA) activity in pAS30
XX or pAS32.
XX KM Enzyme; cancer therapy; leukemia therapy; ss.
XX OS Erwinia chrysanthemi.
XX FH Key Location/Qualifiers
FT misc_feature 1..25
FT /tag= a
FT /label= pUC9 linker
FT sig_peptide 61..723
FT /tag= b
FT mat_peptide 724..1707
FT /tag= c
FT /note= "Cloned"
FT terminator 172..1789
FT /tag= d
XX PN EP211639-A.
XX XX 25-FEB-1987.
XX PD 04-AUG-1986; 86BP-00305984.
XX PF 06-AUG-1985; 85GB-00019753.
XX PR (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
XX PA Atkinson A, Minton NP, Gilbert HJ;
XX PI WPI: 1987-051784/08.
XX DR P-PSDB; MAP70348.
XX PT New recombinant plasmids coding for L-asparaginase - esp. from Erwinia
XX carotovora, useful for treating malignancies and new transformed hosts.
XX PS Disclosure; Fig 4; 35pp; English.
XX CC When E. carotovora SCRI 193 contg. plasmid pAS32 was cultured, the cell-
XX free culture broth contained 11.19 LA units/ml with specific activity
XX 53.34 units/mg of protein. These figures compare with 10.87 LA units/ml
XX and 14.97 units/mg of protein for the currently used prodn. strain
XX E. chrysanthemi NCPB 1066. (Updated on 25-MAR-2003 to correct PA field.)
XX XX
SQ Sequence 2837 BP; 627 A; 665 C; 867 G; 678 T; 0 U; 0 Other;

Query Match 25.8%; Score 261.6; DB 1; Length 2837;
Best Local Similarity 55.6%; Pred. No. 2.7e-36;
Matches 554; Conservative 0; Mismatches 424; Indels 18; Gaps 2;

QY 20 AGAAGTGGCCAAAGTGTATCTGTGGCCACCGCGGACCAATGCGCGGCTGCGGCCCA 79
DB 728 ATAAAGTCCCAATATGTTATCTGTGCGACCGCGGTACAAATGCGCGCTCAAGCGCAA 787
QY 80 GCGCGGCAACAGCGCCACCTACAGGCTGCGCAAGTTGGGTGCTGACAGCTGATGCGG 139
DB 788 CGGCTTACCAACCAACGATTACAGGCTGCGCGCTTGGGTGATACGTTAATCAACG 847
QY 140 GCGTCCGAGACTGGCGCACTGCGCAATGTGCGCGCGCGACAGGTGATGACGATCGCT 199
DB 848 CTGTGCTGAGGTGAAGAAACTGTGCTATGTGAAGGGGAGCACTTCTCCAAATGCGCA 907
QY 200 CCGAAGCATCAACCAAGCAGCATGCTGTCAAGCTGGCAAGCAAGCGTGGCGGCGG 259

```

[illegible]

XX	01-JUN-1998;	98CN-00102046.
PF		
XX	01-JUN-1998;	98CN-00102046.
PR		
XX	(MICR-) INST MICROBIOLOGY CHINESE ACAD SCI.	
PA		
XX	Qian S, Wang Y, Meng G;	
XX	WPI; 2000-351194/31.	
DR	P-PSDB; AAM90867.	
XX		
PT	New structured L-asparaginase bacterial host cell and its production	
PT	culture.	
XX		
PS	Disclosure; Fig 2, 12pp; Chinese.	
XX		
CC	This invention describes a novel L-asparaginase-producing recombinant	
CC	bacteria and its fermentation culture method. The new cell is produced	
CC	using standard recombinant techniques. The enzyme gene expressed in the	
CC	cell is different from the reported enzyme gene in sequence, and its	
CC	individual amino acid composition is also different. The culture medium	
CC	adapted for culturing the new bacterial cells is cheap and easily	
CC	available, and the fermentation process used to grow the cells and	
CC	express the enzyme is simple and easy to implement. This sequence encodes	
CC	an L-asparaginase which is described in the method of the invention	
XX		
Q0	Sequence 1044 BP; 274 A; 281 C; 272 G; 217 T; 0 U; 0 Other;	

Query Match	24.9%	Score 252.2	DB 3	Length 1044
Best Local Similarity	55.7%	Pred. No. 1.2e-34		
Matches 548	Conservative 0	Mismatches 423	Indels 12	Gaps 3

[illegible]

Db 1643 CCGCCGCGAAAACGGGTACTGTCAGTCGTGGCTTCCCGCGTACCGAGCGGGCGCTAACGA 1702
 QY 866 TCGGTAAACCCCGAGCGACCCCGAGACAAAGACGACTGGGTCTGGTCCCGACGACTTGAAC 925
 Db 1703 CTCAGATCCCGAAGTGAATGATGCGAAATACGGCTTCGTGGCTTGGCAACGCTGAAC 1762
 QY 926 CCGAGAAAGCCCGCATCTCTGCGATGTGGCAATGACCAAGCCAGAGCAGCAAGAGC 985
 Db 1763 CCGAAAAACCGCGCTTCTGTCTCAACTGGCTCTGACGCAACCAAGATCCGACAGCA 1822
 QY 986 TCGACGCGATTTTCTGGGAATCTGA 1011
 Db 1823 TCCAGCAGATCTTCATCACTACTAA 1848

RESULT 10

ACH94935
 ID ACH94935 standard, DNA, 1071 BP.

AC ACH94935,
 XX
 XX

DT 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polynucleotide seqid 730.

XX Recombinant expression vector; transcription regulatory element;
 KM Klebsiella pneumoniae protein; antibacterial; vaccine; gene; ds.
 XX Klebsiella pneumoniae.

OS US610836-B1.
 XX
 XX

PN 26-AUG-2003.
 XX
 XX

PD 27-JAN-2000; 2000US-00489039.
 XX
 XX

PR 29-JAN-1999; 99US-0117747P.
 XX
 XX

PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 XX

PI Breton GL, Osborne M,
 XX
 XX

DR WPI, 2003-895346/82.
 XX
 XX

DR P-P8DB; ABO61384.
 XX
 XX

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.

XX Disclosure; SEQ ID NO 730; 932bp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
 CC pneumoniae polypeptide of the invention
 CC
 XX

XX Sequence 1071 BP; 229 A; 281 C; 338 G; 223 T; 0 U; 0 Other;

XX Query Match 23.7%; Score 240.4; DB 11; Length 1071;
 XX Best Local Similarity 54.2%; Pred. No. 1.3e-32;
 XX Matches 543; Conservative 0; Mismatches 441; Indels 18; Gaps 2;

QY 13 AACGACAGAGCTGGCCAGCTGATCTGGCCACCGCGGCACTCGCCGCGCT 72
 Db 85 ACGAGAGACCGCTTCCGATATTGTTATTTTAAAGCAACCGGAGACGATGCTGTTG 144

QY 73 GCGCGCAGCGCGCCCAACGCGGCTTACAGGCTGCCAAGTTGGCTGACAAAGCTG 132
 Db 145 GCGGCGACGAAATACCAACCAACCGGCTTAAAGCCGGGCGATTGGGCTGACAGCGCTG 204

QY 133 ATTCGCGGCGTCCCGAGCTGGCCGACCTTGGCAATGTCGCGGAGAGAGGTATGACG 192
 Db 205 ATTAAAGCCGTCCCGAATATGAGTAAATGCTCAAGTGAAGGCGAGAGAGGTGCAAT 264
 QY 193 ATTCGCTCCGAAAGCATCAACGACGACCTGTCTAGCTGGCAAGAGCTGGCCGAG 252
 Db 265 ATTTGGCAGTGAATATGACCAAGCATATATTTCTCAAGCTCTCGAAGCGGGTGAATGCG 324
 QY 253 CTGGCCGACAGATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 312
 Db 325 CTATTGGCTCGGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 384
 QY 313 GAAACCGCTTACTTTTGAACCTCGTGAAGAAAGACGCAAGACCGATCTGTGTGCGGT 372
 Db 385 GAAACCCGTATTTCTCTCAATCTGACGATGAAAGAAAGCAACCGGTGATCTTTACCGCG 444
 QY 373 TCCATGCGCCCGCGACCGCCATGTCGCGCGACCGGATGCTCAACCTGTACACGCGGTG 432
 Db 445 GCGATGCGCCCGCGACCGCCATGTCAGCGCGCGACCGGATGACCTGTGGAAGCGGTG 504
 QY 433 GCGGTGCGGACGACAGACGATCTGCGCGGCAAGGCGGTGTGACCAATGAAACGACGAG 492
 Db 505 ACGGTGCGCGCGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 564
 QY 493 ATTCAGTCCGCGCGTGAACGTCGACCAAGTGAATCAATCAAGACCGAAGCTTCAAGAGC 552
 Db 565 ATTCGCGCGCGCGCGCTTTGTACCAAACTAATGCACTTGTGTGATCTTTCCGGGCG 624
 QY 553 GCGTGGG--GCCGCTGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609
 Db 625 CCGGAGAGGCGCTATCTTGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGT 684
 QY 610 GCCAAGCGGACAGGCTCACTCCGATGTCATCAATCAAGACATCAAGACCTTCCCGCG 669
 Db 685 GATTAATTTCAACGCTGCGCTGCGTGTGATGATGATGATGATGATGATGATGATGATGAT 744
 QY 670 GTGACATGCGCTTACAGTATGACCAAGCTCAACGACGACGCTTACCAAGGCTTGGACAG 729
 Db 745 GTGTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 804
 QY 730 AACGCGCGCAAGCGCTGATTCATGCTCCGACCGGCAATGCTGCTGCTGCTGCTGCTGCT 789
 Db 805 CATCATGCGGACGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 864
 QY 790 GTGCGAGCGCTGACAGAGCTGCGCAAGACGCGGTGCAATATTTGCTTGTGCAAGTCA 849
 Db 865 GCGCGCGGTATMAAAAGCGGACGAGCGGCAATGTCGTGTGTGTGTGTGTGTGTGTGTGT 924
 QY 850 CAGGCGGCTTGTGTCGTGCTGTAACGCGCGGACCGGCAAGCAAGACGATGGGTGCTG 909
 Db 925 GCGAGCGGCTGTGTCGTGCTGTAACGCGCGGATGACGCGCGG-----GTTGGTG 969
 QY 910 GCCACGACCTTGAACCGGCAAGGCGCGCATCTGGCGATGTGGCAATGACCAAGACC 969
 Db 970 GCCGACTCCCTCAACCGGCAAGGCGGATATGTCGATGACGCGGCTGACGAGACT 1029
 QY 970 CAGGACAGAGAGCTGACGCGCATTTTCTGGGAATCTGA 1011
 Db 1030 AAGATCCGACGCTGATTCAGCAATATTTCTACTACTGA 1071

RESULT 11

AAK03474
 ID AAK03474 standard, DNA, 1133 BP.

XX AAK03474;
 XX
 XX

DT 30-APR-1999 (first entry)

XX Wolinella succinogenes L-asparaginase DNA.

XX L-asparaginase; amplification; treatment; disease; asparagine depletion;
 KW malignant disease; haematology; lymphoma; leukaemia; myeloma; AIDS;

KW autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
 KW covalent modification; acylation; pharmacokinetic; immunogenic; spleen;
 KW hypersensitive; reduced toxicity; immunosuppression; allergy; thymus;
 KW lymphocyte marker distribution; hepatotoxicity; ss.

XX Molinella succinogenes.

XX WO9856410-A1.

XX 17-DEC-1998.

XX 09-JUN-1998; 98WO-US011905.

XX 09-JUN-1997; 97US-0049085P.

XX (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.

XX Durden DL;

XX WPI; 1999-070295/06.

XX Treating diseases responsive to asparagine depletion - especially cancer
 PT and auto-immune diseases, using a Molinella succinogenes asparaginase.

XX Example 5; Fig 6; 58pp; English.

XX This sequence is a novel L-asparaginase gene from Molinella succinogenes
 CC which is used in a method for the treatment of diseases that respond to
 CC asparagine depletion. This asparaginase and its analogues, are
 CC particularly used to treat (i) malignant diseases, particularly where
 CC haematological and specifically lymphoma, leukaemia and myeloma (in
 CC either chronic or acute phases) and (ii) autoimmune diseases,
 CC specifically rheumatoid arthritis, systemic lupus erythematosus and
 CC acquired immune deficiency syndrome, in humans or other mammals. The
 CC asparaginase can also be used for the covalent modification of
 CC biologically active proteins by acylation which is used to alter
 CC pharmacokinetic and immunogenic properties of biologically active
 CC proteins, and can be applied to any protein used to treat animal, human
 CC or plant diseases. The asparaginase can be used to treat patients who
 CC have become hypersensitive to other microbial asparaginases. The
 CC asparaginase has reduced toxicity and reduced tendency to cause
 CC immunosuppression or allergy, particularly it has no effect on spleen and
 CC thymus histology or lymphocyte marker distribution and no hepatotoxicity
 CC
 SQ Sequence 1133 BP; 315 A; 305 C; 267 G; 246 T; 0 U; 0 Other;

Query Match 21.6%; Score 219; DB 2; Length 1133;

Best Local Similarity 52.8%; Pred. No. 6,4e-29;

Matches 521; Conservative 0; Mismatches 460; Indels 6; Gaps 2;

QY 29 CCNAACGTGTGATCTCTGCGCACCGGCGGACCATCCCGCGCTGCGCGCAGCGCGGCA 88
 DB 128 CCCAAGTGAATCTTCTAGCAGACGAGGACCATCTGCTGCGGCGAATCTAGGCTCA 187
 QY 89 ACAGCGCCACTTACCAAGGCTGCCAAGGTTGGGTGCAAGCTGATTCGCGGCGTCCG 148
 DB 188 AAGAG--TAGTACTCTGCTGAGGAGTCAACGCTTATTAAGCTTCTTGACGCGTCCCTG 244
 QY 149 AGCTGCGCGACCTGCGGCATATGCGCGGAGAGAGTATGCAAGTCGCTCCGAAAGA 208
 DB 245 CCATCAACCACTTACGACCATCAAGGAGTGAAGATTCACAGCATTTGCTCCCAAGAGA 304
 QY 209 TCACCAACGACGACCTGCTCAAGCTGCGCAAGCAGCTGCGCGAGCTGCGCAGCAATG 268
 DB 305 TGAACGGTAAAGTGTGCTTAACTAGCCAGGCTGTCAATAGCTCTCTCCGCAAAAG 364
 QY 269 AGGTGATGATCTGCTATCAATCCATGCAACCGACCTCTGAAAGAACCGCTTACTTT 328
 DB 365 AAGACGAAACCGTATCATCAACCATGAACTGACACCATGAAAGAGACCGCTTCTTCC 424
 QY 329 TGAACCTGCTGAAAAAGACGACCAAGCGATGCTGTGCTGCTTCCATGCGCCCGGCA 388
 DB 425 TCAACCTCAAGGTGAAGAACCAAAACCTGTGCTTGTAGCGCGCATGCTTCAGGCT 484

QY 389 CGGCATGTCGCCGACGCGATGCTCAACCTGTACAAAGCGGTGCGCTGCGCAGCAACA 448
 DB 485 CTTTCATAGTGTGATGCGCCCATGAATCTCTAATACGCGGTGAATAGCATCAACA 544
 QY 449 AGAATCGCGCGGCGAAGGCGCTGCTGTGACATGAAGAAGATGCAATGATCCGCGGCTG 508
 DB 545 AAGCTTACTTAAACAAAGAGTGTGATTTGTGAAGATGATGATTCACGCGCAAG 604
 QY 509 AGGTGACGATGATCAATCAACATCAAGACCGAGCTTCA--GAGCGCTGCGGCGCCGC 565
 DB 605 AAGGACCAAGCTCAACACACACCGAGTCAATGATTTGCTTGGCCCAACAGGTAAAG 664
 QY 566 TGGGATGTTGATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 625
 DB 665 TCGGACAGTCTAATATGAGCAAGTCGATTTTCACTCAATCCGTTGACCTCACACC 724
 QY 626 TCAACTCGGATTTGACATCAAGCAAGATACAGACGCTGCGCGGATGCAATCGCTTCA 685
 DB 725 TTGCAAGTGAATTTGATTTATGCAAAATGCAAGAACTCCAGAGTCAATTTCTTACG 784
 QY 686 GCTATGCAAGCTCAACGACGCGCTTCAAGGCTTGGCAGACAGACGCGCAGAGGCGC 745
 DB 785 CTCACCCGATGATGATGATGATTTTATGATGATGATGATGATGATGATGATGATGAT 844
 QY 746 TGAATCATGCGCGGACCGGCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 805
 DB 845 TCATCATGACGAGCATGGGCAATGGGAAACCTTTCCCTTTCATCAAAATGCTTTGAAA 904
 QY 806 AGCTGCGCAAGACGCGCTGCAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 865
 DB 905 AAGCAGCCAAATCAAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 964
 QY 866 TGGCTAAGCGCGAGGACCGGACGACGACGACGACGACGACGACGACGACGACGACG 925
 DB 965 CCCAAGAGGCTGAAGTGAATGATGAAGAACTTGTGTTGTGCTGCTGCTGCTGCTGCTG 1024
 QY 926 CGCAGAGGCGCGCATCTGCGGATGATGATGATGATGATGATGATGATGATGATGATG 985
 DB 1025 CTCAAAAGGACGAGTGTCTTCTTATGTTAGCCCTCACCAAACTGATGATGAGAGGCA 1084
 QY 986 TGCAGCGCATTTCTGGAATATGAT 1012
 DB 1085 TCCAAAAGATCTCTCACTTATPAT 1111

RESULT 12

AAC62512 ID AAC62512 standard; DNA; 1133 BP.

XX AAC62512;

XX 07-FEB-2001 (first entry)

XX W. succinogenes asparaginase-specific DNA insert.

KW Asparaginase; glutaminase; autoimmune disease; graft versus host disease;
 KW leukaemia; DNA insert; de.

XX Molinella succinogenes.

XX WO200059533-A1.

XX 12-OCT-2000.

XX 24-MAR-2000; 2000WO-US007981.

XX 02-APR-1999; 99US-0127662P.

XX (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.

XX Durden DL;

DR WPI; 2000-664963/64.

XX Treating or preventing graft versus host and autoimmune diseases, e.g.
PT diabetes, rheumatoid arthritis, that responds to asparagine or glutamine
PT depletion comprises administering an asparaginase or glutaminase.

PS Example 5; Fig 6; 72pp; English.

XX The present invention relates to a novel method of using the recombinant
CC microbial enzymes asparaginase and glutaminase to treat autoimmune
CC diseases and leukemia. The diseases likely to respond to such treatment
CC include graft versus host disease, rheumatoid arthritis, systemic lupus
CC erythematosus, autoimmune myelitis, collagen vascular diseases, AIDS,
CC osteoarthritis, Isaac's syndrome, psoriasis, diabetes, multiple
CC sclerosis, sclerosing pancreatitis, rheumatic fever, inflammatory
CC bowel diseases, primary biliary cirrhosis, chronic active hepatitis,
CC glomerulonephritis, myasthenia gravis, pemphigus vulgaris and Graves
CC disease. The present sequence is DNA insert used to demonstrate the
CC method of the invention

80 Sequence 1133 BP, 315 A; 305 C; 267 G; 246 T; 0 U; 0 Other;

Query Match 21.6%; Score 219; DB 3; Length 1133;

Best Local Similarity 52.8%; Pred. No. 6.4e-29;

Matches 521; Conservative 0; Mismatches 460; Indels 6; Gaps 2;

QY 29 CCAAGCTGTGATCTCTGGCCACCGGCGACCATTCGCGCGCTGGCGCGCGCGCA 88
DB 128 CCCAAGTGTACTTCTTACGACGAGGAGGACCATTCCTGCTTGGGGAATCTAGCGTCA 187
QY 89 ACAGCGCCACTTACAGGCTGGCGGCGTGTGACAGATGATTTGCGGCGCGCG 148
DB 188 AAGG---TAGTACTCTGCTGAGAGCACTACCTTATAGCTTCTTGAAGCGCTCC 244
QY 149 AGCTGGCCGACCTGGCCAAATGTGCGCGGAGAGGTGATGCAATGCGCTCGAAGA 208
DB 245 CCATCAACGACCTTACGACCATCAAGGTGAGAACATCTCAAGCATTTGGCTCCAGAG 304
QY 209 TCACCAACGACCTTCTCAAGCTGGCGAGCGCTGGCGCGAGCTGGCGCGAGCAAG 268
DB 305 TACCGGTGATGAGGTGCTTAACTAGCCAGGTCTCAATGAGCTCTCGCCCAAAAG 364
QY 269 AGCTGATGATCTCTGACATCCATGGACCGACACCTTGAAGAAACCGCTTACTTT 328
DB 365 AAGACGACCTGATCATCACTGATGATCACTGATGATGATGATGATGATGATGAT 424
QY 329 TGAACCTGCTGAAAAAGACCGACAGCCGATCTGCTGCTGCTGCTGCTGCTGCTG 388
DB 425 TCAACCTGACGGTGAAGAAAGCAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 484
QY 389 CCGCCATCTCCGCGACCGGATCTCACTGATCAACCGCTGCGCTGGCGCGAGCA 448
DB 485 CTTCATGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544
QY 449 AGGACTCGGCGCGCAAGGCGTGTGATGATGATGATGATGATGATGATGATGATG 508
DB 545 AAGCTCTTACTAAGAGAGGTGATGATGATGATGATGATGATGATGATGATGATG 604
QY 509 AGGTGACAGTGTGATCAATCAAGACCGAACCTTCAA--GAGCGCTGGGCGCGC 565
DB 605 AAGCGACCAAGCTCAACACCGAGTCAATGATGATGATGATGATGATGATGATGAT 664
QY 566 TGGGCTGCTGATGAGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 625
DB 665 TGGGCAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 724
QY 626 TGAACCTCGAGTGTGATCAATCAAGACCGAACCTTCAA--GAGCGCTGGGCGCGC 685
DB 725 TTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
QY 686 GGTATGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 745
DB 785 CTCACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 844

QY 746 TGATCATCGCGCACCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 805
DB 845 TATCATCATGAGGAGTGGGCAATGGAACCTTTCCCTTGAAGTCAAAATGCTTTGAA 904
QY 806 AGCTGGCAAGAGCGGCTGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 865
DB 905 AAGCAGCCAAATCAGGCGGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964
QY 866 TGGCTAGCCGAGACGCGGACGACGACGACGACGACGACGACGACGACGACG 925
DB 965 CCCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1024
QY 926 CGCAGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 985
DB 1025 CTCAGAGCCGAGAGTCTTCTTATGCTTGAAGCTTCAACCAAACTAGTATAGAGG 1084
QY 986 TGCAGGCAATTTCTGGAATCTGAT 1012
DB 1085 TCCAAAGATCTTCTCACCTATTAAT 1111

RESULT 13

ABZ80801

ID ABZ80801 standard; DNA; 1174 BP.

XX

XX ABZ80801;

XX

DT 27-OCT-2003 (revised)

DT 30-MAY-2003 (first entry)

XX

DE Erwinia carotovora L-asparaginase gene.

XX

KW gene; de; asparaginase; cytostatic; blood; acute lymphoblastic leukemia.

XX

OS Pectobacterium carotovorum.

XX

FH Key Location/Qualifiers

FT CDS 48..1094

FT /tag= a

FT /product= "L-asparaginase"

FT /EC_number= "3.5.1.1"

FT primer_bind 48..72

FT /tag= b

FT /bound_molecly= "primer N204"

FT 1078..1100

FT /tag= c

FT /bound_molecly= "primer C204"

XX

PN MO2003018742-A2.

XX

PD 06-MAR-2003.

XX

PF 21-AUG-2002; 2002WO-RU000405.

XX

PR 22-AUG-2001; 2001RU-00123442.

XX

PR 04-APR-2002; 2002RU-00108505.

XX

PA (UCHR=) GOS UCHREZHENIE NI INST BIOMED KHIMI.

XX

PA (UCHR=) GOS UCHREZHENIE TSENTR BIOINZHENERIYA.

XX

PI Eldarov MA, Zhgun AA, Geraziev YV, Aleksandrova SS, Bogush VG;

PI Sidoruk KV, Sveshnikova EV, Borisova AA, Omeinyuk NV, Archakov AI;

PI Skryabin KG, Sokolov NN;

PI

DR WPI; 2003-268418/26.

DR P-PSDB; ABP98615.

XX

XX New isolated Erwinia carotovora L-asparaginase gene useful for producing

PT a L-asparaginase protein useful for treating malignant and nonmalignant

PT diseases of the blood.

XX

PS Claim 2; Page 22-23; 29pp; Russian.

XX This sequence represents a DNA fragment coding for an *Erwinia carotovora*
 CC L-asparaginase. The invention also includes a homogeneous preparation of
 CC a mature form of L-asparaginase produced by expression of the DNA in a
 CC bacterial strain and has low affinity for L-glutamine. The L-asparaginase
 CC protein encoded by the DNA is useful for treating malignant and non-
 CC malignant diseases of the blood (e.g. acute lymphoblastic leukemia).
 CC (Updated on 27-Oct-2003 to standardise OS field)

XX Sequence 1174 BP; 290 A; 284 C; 313 G; 287 T; 0 U; 0 Other;

Query Match 20.6%; Score 209.2; DB 8; Length 1174;
 Best Local Similarity 54.1%; Pred. No. 3.2e-27;
 Matches 471; Conservative 0; Mismatches 393; Indels 6; Gaps 2;

```

20  AAGAGCTGGGCAACGCTGATCTCGGACCGGCGGCACTCGCGGCGGCGGCA 79
Db  118  AAAAACTACTTACCTATTCATTCTGCGACAGCGCGTTCGATTGCCGAGCA 177
Qy  80  GCGCGGCAACAGCGGCACTTACAGGCTGCCAGGTTGGCGTGCACAGCTGATTGCGG 139
Db  178  GCGAATAGCAACCACTGGTAAAGCGGGTGGCGCTGGCGGTAAAGACGCTGATTCAAG 237
Qy  140  GCGTGGCGGAGCTGGCGGCACTGGCGCAATGTGCGGCGGCAAGGTTGATGCAATGCGCT 199
Db  238  CGGTGCTGAGCTGAAAGCGCTTGCCTTCAAGGTGAGAGGTTGCGACATCGGCA 297
Qy  200  CCGAAGCATCAACAGCAAGCACTGCTCAAGCTGGCAAGACGCGGCGGAGCTGCGCG 259
Db  298  GTGAAATATGACAGCGCATGTGTTTAACTGTAAGCAAGCGCGTGAACAGAGCTGCTG 357
Qy  260  ACAGCAATGACGTGATGCGATCGTCAATGACCCATGACCGCAACCTGGAAGAAACCG 319
Db  358  CAGCGACGATGTTATGCGTGTGATCAACGAGTACGATACGCTGCAAGATCTC 417
Qy  320  CCTACTTTTGAACCTCGTGGAAAAAGACGCAAGCCATCGTGGTGGCTGCTGATGCG 379
Db  418  CTTATTTCTGAACTGAGCGGTGAAAGGACAGCAACGATACCTTTTGCATGCGTCCG 477
Qy  380  GCCCGGCAACCGCATGTCCGCGACGCGGATCTCAACTGTACAAAGCGCGTGGCGTGG 439
Db  478  GCAAGGCGCATCAAGT---CCGAGCGGCCCATGAACTGTATCGGTGCAATGAAAGTGG 534
Qy  440  CCAGCAACAAAGACTCGCGCGGCAAGGCGGCTGTGTGCAATGAAAGCAAGATCCAGT 499
Db  535  CAGCGGATGAAATCTCCGCGGTGCGGTGTACTGGTGTGCTGAACGACCGCATTTGGTT 594
Qy  500  CCGGCGGTGACGTGAGCAAGTGCATCAATCAAGACCGAAGCTTTCA---GAGCGCT 556
Db  595  CTGCCCGTTTCACTGACCAAAACCAAGCCTTTACGTTGATACCTTTAAAGGCCAGAG 654
Qy  557  GGGGCGCGGTGAGTGTGTGAAAGCAAGTGTGTTCCGCTGCGCGGCAAGC 616
Db  655  AAGGTTATCTGGGCGGTGATTAATCGGTGACAAATATCTACAGACCGCTGTGATTAAG 714
Qy  617  GCCACACGCTCACTCCGAGTTTGCACATCAAGCAATGCAAGCTCCGCGGAGTGAACA 676
Db  715  TTCAACACCGGCTTCGTTGTTGATGTAACCAAGTTATGCTGCGGAAAGTGCACA 774
Qy  677  TCGCTTACAGCTATGAGCACTGCACGAGACCGGCTTACAGGCGCTTGGACAGAAACGCG 736
Db  775  TTATTTATGTTATATCAGACGATCCAGAAATATATGATATACCTCTTATCAAAACGCGG 834
Qy  737  CCAAGGCGGTATCATATGCGGCAACCGGCAATGCTCGGTGTGCTGCGGGTGTGCGCA 796
Db  835  TAAAGGATCTGTATGCGGCGATGCGGCGACGCTATTCAGAGCGCGGCGACGCTG 894
Qy  797  CCTGTCAAGAGTGTGCGCAAGAACGCGGTGCAATCTGTTGTTCACTCAACAGGCGG 856
Db  895  GCATCCGTAAAGCGGAAACAAAGGCAATGTTGTGTAAGTTCAGACGCTACCGGCAAGG 954
Qy  857  GTTTCGTGTGCTAAGCGCGGACGCGCG 886

```

Db 955 GTATGCTCCACCGGATGACAGGCCAACCCG 984

RESULT 14
 ABD07233/C
 ID ABD07233 standard; DNA, 270 BP.

XX ABD07233;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polynucleotide #5837.

KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
 KW antibacterial.

OS Pseudomonas aeruginosa.

FN US651195-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR P-PsDB; AB073662.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnosis, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 5837; 455bp; English.

The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnosis,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biophysics technology. Sequences ABD01397-
 CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html

XX Sequence 270 BP; 39 A; 92 C; 85 G; 54 T; 0 U; 0 Other;

Query Match 20.4%; Score 207.2; DB 11; Length 270;
 Best Local Similarity 85.8%; Pred. No. 7.2e-27;
 Matches 230; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```

321  TACTTTTGAACCTCGTGAAGAAAGCAGCAAGCCGATCGTGTGCTTCATCG 380
Db  270  CTACTCTCTGACCTGTGTGAGACACCGAAGAGCTTATCGTGTGCTGCTCATGCG 211
Qy  391  CCGCGGACCGGCAATGTCGCGGCAACGCGATCTCAACTGTACAAAGCGGTGGCGGTGCG 440
Db  210  CCGGCGACCGCGATGTGCGCGGACGCGATCTCAACTGTACAAAGCGCGGTGGCGGTGCG 151
Qy  441  CAGCAACAGAGACTGCGCGGCAAGGCGGTGTGTGATCAATGAACAGAGATCAAGTGC 500

```

DB 150 CGCGCAAGATCGGCAAGCGGCGTGTATCACCATTGAACGACGAGATCTCTTC 91
 QY 501 CGGCGGTGACGTGAGCAAGTTCATCAATCAAGACCGAAGCTTGAAGAGCGCCCTGGGG 560
 DB 90 CGGCGCGGACGCGAGCAAGATGTGTCAACATCAAGACCGAAGCGTTCAGAGCGCCGCGGG 31
 QY 561 CCCGCTGGGACATGCTGTGGAAGCAAG 588
 DB 30 CCGGCTGGGACATGCTGTGGAAGCAAG 3

RESULT 15

AD007178
 ID AD007178 standard; cDNA, 1303 BP.

XX AD007178;

DT 15-JUL-2004 (first entry)

DE Aspergillus oryzae asparaginase coding sequence.

XX Asparaginase; enzyme; acrylamide; heat treatment; gene; ss.

XX Aspergillus oryzae.

PH Key Location/Qualifiers

FT CDS 49..1185

FT /tag a

FT /product= "Asparaginase"

FT /EC_number= "3.5.1.1"

XX MO2004032648-A1.

XX 22-APR-2004.

XX 10-OCT-2003; 2003WO-DK000684.

XX 11-OCT-2002; 2002DK-00001547.

XX (NOVO) NOVOZYMES AS.

PI Budolfesen G, Jensen MT, Heide-Hansen HP, Stringer MA, Lange L;

DR P-PSDB; AD007179.

XX Claim 11; SEQ ID NO 1; 46bp; English.

XX The present sequence is that of cDNA encoding Aspergillus oryzae
 CC asparaginase. The cDNA was isolated from a cDNA library on the basis of
 CC homology to Saccharomyces cerevisiae asparaginase II. The enzyme shows
 CC optimal activity at pH 7 (at 37 degrees C) and 60 degrees C (pH 7.0) and
 CC is stable in the pH range from about 4.0 to 8.0. It has a molecular
 CC weight of 40-65 kDa (SDS-PAGE). A claimed method of reducing the
 CC formation of acrylamide during heat treatment of raw material comprising
 CC carbohydrate, protein and water involves treatment of the raw material
 CC with an enzyme prior to heat treatment. The enzyme may be an
 CC asparaginase, including A. oryzae asparaginase, that is capable of
 CC reacting on asparagine or glutamine (optionally substituted). The method
 CC is useful for preparing a heat-treated product (claimed) such as potato
 CC products, potato chips, breakfast cereals, biscuits, roasted nuts, hot
 CC cakes, pancakes, etc.

SQ Sequence 1303 BP, 303 A; 421 C; 305 G; 274 T; 0 U; 0 Other;

Query Match 20.3%; Score 205.6; DB 12; Length 1303;
 Best Local Similarity 56.5%; Pred. No. 1.3e-26;
 Matches 445; Conservative 0; Mismatches 334; Indels 9; Gaps 3;

QY 25 CTGGCAACGTGATCTTGGCAACCGGCGGACCATGCGCGGCTGGCGCAAGCGG 84
 DB 199 CTGGCAACGTGATCTTGGCAACCGGCGGACCATGCGCGGCTGGCGCAAGCGG 258
 QY 85 GCCAACAAGCGGCACTTACCAAGGCTGCCAAGTTGGGTGACCAAGCTGAATTCGCGGTG 144
 DB 259 ACGGCAACAAGGTTACCAAGGCTGCCAAGTTGGGTGACCAAGCTGAATTCGCGGTG 318
 QY 145 CCGGAGCTGGCGGACCTGGCAATGTGCGCGGCGGAGAGGTGATGCAAGTCCGCTCCGA 204
 DB 319 CCGGAATGCTAAAGGTTCCCAAGCTGCTGGCGGTGCAAGTAACTCCGAGCCCA 378
 QY 205 AGCATCAACCAAGCAAGACCTGCTCAAGCTGGAAGAGGCTGGCGGCGGCAAGC 264
 DB 379 GACATCACTTCCGCAATTTCTCTGCGCTCTTCAACAGATCAAGAGGTGCTGCAAC 438
 QY 265 AA---TGAAGTCAATGCGATCGTCAATCAACCAATGCAACCGAACCCTGGAGAAAACGCGC 321
 DB 439 GACCCCAACATGCGCGGTGCAATGATGATCAACCGGCAACGCTGCAAGAAATCGCGC 498
 QY 322 TACTTTTGAACCTGCTGGAAGAAAGACCAAGCGGATGCTGCTGCTTCCATGCGC 381
 DB 499 TTCTTCTCGACGCCACGCTCAATGTCGCAAGCCCGTGTGATGCTGCGCGCATGCGC 558
 QY 382 CCGGCAACCGGATGCTGCGCGGCAAGCGGCAAGCTCAACCTGTAACAACGCGGCGGCGC 441
 DB 559 CTTTCAACCGGCACTTGGGTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 618
 QY 442 AGCAACAAGGACTGCGCGGCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
 DB 619 GCGAGCCCAAGGCGGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 678
 QY 502 GCGCGTGAAGTGAAGTGAATCAATCAATCAAGCAAGCAAGCAAGCAAGCAAGCAAG 558
 DB 679 GCGTTCAGCGCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 738
 QY 559 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 618
 DB 739 GGTAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 795
 QY 619 CACACGTCACCTCCAGTTCGACATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 678
 DB 796 CCAACAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 855
 QY 679 GCGTACAGCTATGCAAGCTCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 738
 DB 856 CTCTACTCATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 915
 QY 739 AAGCGCTGATCAATGCGGCGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 798
 DB 916 AAGGCGATGTTATGCGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 975
 QY 799 CTGACAGA 806
 DB 976 ATGGAAGA 983

Search completed: March 22, 2005, 10:31:22
 Job time : 664 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: March 22, 2005, 07:01:33 ; Search time 216 Seconds
(without alignments)
7681.410 Million cell updates/sec

Title: US-09-842-628-1

Perfect score: 1014

Sequence: 1 aaggaagcgcgagaccagca.....ttctctggaatcgtatga 1014

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/6C_COMB.seq:*
6: /cgn2_6/prodata/1/ina/6D_COMB.seq:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1014	100.0	1014	5	PCT-US92-10421-1
2	983.4	97.0	1017	3	US-08-050-482A-1
3	742	73.2	1182	4	US-09-252-991A-5793
4	742	73.2	1335	4	US-09-252-991A-5877
5	562.2	55.4	2748	4	US-09-252-991A-5773
6	273	26.9	1125	4	US-09-328-352-2336
7	240.8	23.7	1848	1	US-08-447-422-15
8	240.4	23.7	1071	1	US-09-489-039A-730
9	219	21.6	1133	3	US-09-094-435-3
10	207.2	20.4	270	4	US-09-252-991A-5837
11	138	13.6	1830121	4	US-09-557-884-1
12	138	13.6	1830121	4	US-09-543-990A-1
13	129	12.7	1041	4	US-09-543-681A-1219
14	83	8.2	4403765	3	US-09-103-840A-2
15	83	8.2	4411529	3	US-09-103-840A-1
16	78.8	7.8	2272	4	US-09-252-991A-5838
17	72.2	7.1	963	4	US-09-583-110-925
18	72.2	7.1	966	4	US-09-107-433-791
19	66.8	6.6	1005	4	US-09-107-532A-1293
20	66.2	6.6	1002	4	US-09-134-000C-1208
21	65	6.4	1104	4	US-09-252-991A-9014
22	65	6.4	1425	4	US-09-252-991A-8780
23	65	6.4	1653	4	US-09-252-991A-10546
24	65	6.4	1701	4	US-09-252-991A-10316
25	64.6	6.4	1758	4	US-09-252-991A-8997
26	64.6	6.4	3120	4	US-09-252-991A-6505
27	64.6	6.4	3122	4	US-09-902-540-469

28	64.4	6.4	1268	2	US-08-403-852D-2	Sequence 2, Appl1
29	64.4	6.4	1268	3	US-08-510-646B-2	Sequence 2, Appl1
30	64.4	6.4	1268	3	US-09-231-81B-2	Sequence 2, Appl1
31	64.4	6.4	1268	4	US-09-635-359B-2	Sequence 2, Appl1
32	64.4	6.4	5392	2	US-08-403-852D-1	Sequence 1, Appl1
33	64.4	6.4	5392	3	US-08-510-646B-1	Sequence 1, Appl1
34	64.4	6.4	5392	4	US-09-231-81B-1	Sequence 1, Appl1
35	64.4	6.4	5392	3	US-09-635-359B-1	Sequence 1, Appl1
36	64.2	6.3	876	4	US-09-252-991A-1565	Sequence 1565, Ap
37	64.2	6.3	1512	4	US-09-252-991A-1551	Sequence 1551, Ap
38	64.2	6.3	1677	4	US-09-252-991A-1616	Sequence 1616, Ap
39	62.8	6.2	2810	3	US-09-408-647A-1	Sequence 1, Appl1
40	62.4	6.2	1644	4	US-09-252-991A-792	Sequence 792, App
41	62.4	6.2	2907	4	US-09-252-991A-831	Sequence 831, App
42	62.4	6.2	49225	4	US-09-902-540-1269	Sequence 1269, Ap
43	61.8	6.1	1989	4	US-09-252-991A-15371	Sequence 15371, A
44	61.8	6.1	2058	4	US-09-252-991A-15462	Sequence 15462, A
45	61.8	6.1	2231	3	US-08-961-527-265	Sequence 265, App

ALIGNMENTS

RESULT 1
PCT-US92-10421-1
Sequence 1, Application PC/TUS9210421
GENERAL INFORMATION:
APPLICANT: Robert, Joseph
APPLICANT: MacAllister, Thomas W
APPLICANT: Sethuraman, Natrajan
APPLICANT: Freeman, Abbie G
TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND
TITLE OF INVENTION: ITS USE IN ANTI-VIRAL AND ANTICANCER THERAPY
NUMBER OF INVENTIONS: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Banner, Birch, McKie and Beckett
STREET: 1001 G Street N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10421
CLASSIFICATION:
FILING DATE: 19921204
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: DE P4140003.8
FILING DATE: 04-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 00100.41200
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1014 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: PSEUDOMONAS 7A
PCT-US92-10421-1
Query Match 100.0%; Score 1014; DB 5; Length 1014;

Best Local Similarity 100.0%; Pred. No. 1e-209;
Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AAGGAAGTGGAGAACGAGAGAGCTGGCCAACTGCTGATCTGGCCAGCGCGGCAACC 60
DB 1 AAGGAAGTGGAGAACGAGAGAGCTGGCCAACTGCTGATCTGGCCAGCGCGGCAACC 60
QY 61 ATGCGCGGCGCTGGCGCCAGCGCGCCCAAGCGCCCACTTACAGAGCTGGCCAGGTTGCG 120
DB 61 ATGCGCGGCGCTGGCGCCAGCGCGCCCAAGCGCCCACTTACAGAGCTGGCCAGGTTGCG 120
QY 121 GTGCAAGAGCTGATTTGCGCGCGTGGAGCTGGCCAGCTTGGCCATTTGCGCGCGAG 180
DB 121 GTGCAAGAGCTGATTTGCGCGCGTGGAGCTGGCCAGCTTGGCCATTTGCGCGCGAG 180
QY 181 CAGGTGATGATGATGCTCCGAAAGCATCAACAGAGAGCTGCTCAAGCTGGCAAGC 240
DB 181 CAGGTGATGATGATGCTCCGAAAGCATCAACAGAGAGCTGCTCAAGCTGGCAAGC 240
QY 241 AGCGTGGCCGAGCTGGCCAGAGCAATGACGTGATGATGATGATGATGATGATGATG 300
DB 241 AGCGTGGCCGAGCTGGCCAGAGCAATGACGTGATGATGATGATGATGATGATGATG 300
QY 301 GACACCTCTGGAGAGAAACCGCTTACTTTTGAACCTCTGAGAAAGACCGAACGCGATC 360
DB 301 GACACCTCTGGAGAGAAACCGCTTACTTTTGAACCTCTGAGAAAGACCGAACGCGATC 360
QY 361 GTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
DB 361 GTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
QY 421 TACAAAGCGCGTGGCGCGCCAGCAAGCAAGTCTGCGCGCAAGCGCTGCTGATGAC 480
DB 421 TACAAAGCGCGTGGCGCGCCAGCAAGCAAGTCTGCGCGCAAGCGCTGCTGATGAC 480
QY 481 ATGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 481 ATGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 GCTTCAAGAGCGCTGGCGCGCCAGCAAGCAAGTCTGCGCGCAAGCGCTGCTGATGAC 600
DB 541 GCTTCAAGAGCGCTGGCGCGCCAGCAAGCAAGTCTGCGCGCAAGCGCTGCTGATGAC 600
QY 601 CGCTGCGCGCCAGCGCCAGCAAGCAAGTCTGCGCGCAAGCGCTGCTGATGATGAC 660
DB 601 CGCTGCGCGCCAGCGCCAGCAAGCAAGTCTGCGCGCAAGCGCTGCTGATGATGAC 660
QY 661 CTGCGCGCGAGTGAATGCTGCTTACAGCTTATGAGCAAGTCAAGCAAGCGCTTAC 720
DB 661 CTGCGCGCGAGTGAATGCTGCTTACAGCTTATGAGCAAGTCAAGCAAGCGCTTAC 720
QY 721 CTGCGCAAGAGCGCGCCAGAGCGCTGATCAATGCGCGCAAGCGCGCAATGCTCGG 780
DB 721 CTGCGCAAGAGCGCGCCAGAGCGCTGATCAATGCGCGCAAGCGCGCAATGCTCGG 780
QY 781 TCGCGAGTGTGCTGAGCGCTTACAGAGCTGCGCAAGAGCGCGTGAATGATGCTTGC 840
DB 781 TCGCGAGTGTGCTGAGCGCTTACAGAGCTGCGCAAGAGCGCGTGAATGATGCTTGC 840
QY 841 TCACTGCAAGAGCGCGCTTACAGAGCTGCGCAAGAGCGCGTGAATGATGCTTGC 900
DB 841 TCACTGCAAGAGCGCGCTTACAGAGCTGCGCAAGAGCGCGTGAATGATGCTTGC 900
QY 901 TGGGTGCTGGCCAGAGCTGTAACCGCGAGAGCGCGCGATCTGAGCGATGCTGAG 960
DB 901 TGGGTGCTGGCCAGAGCTGTAACCGCGAGAGCGCGCGATCTGAGCGATGCTGAG 960
QY 961 ACCAAGAGCCAGAGAGAGAGAGCTGCAAGCTTCTGAGAAATGATGATTA 1014
DB 961 ACCAAGAGCCAGAGAGAGAGAGCTGCAAGCTTCTGAGAAATGATGATTA 1014
```

RESULT 2

US-08-050-482A-1
Sequence 1, Application US/08050482A
Patent No. 6312939
GENERAL INFORMATION:
APPLICANT: ROBERTS, Joseph
MACALLISTER, Thomas W.
SETHURAMAN, Nataraajan

TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND ITS USE IN ANTIVIRAL AND ANTICANCER THERAPY

NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington

STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/050,482A
FILING DATE: 25-Apr-1995
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US92/10421
FILING DATE: 04-DEC-1992

ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.

REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 023032/0106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 1017 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS

LOCATION: 1..1011
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-050-482A-1

Query Match 97.0%; Score 983.4; DB 3; Length 1017;
Best Local Similarity 98.6%; Pred. No. 4.1e-203;
Matches 1003; Conservative 0; Mismatches 11; Indels 3; Gaps 1;

```
QY 1 AAGGAAGTGGAGAACGAGAGAGCTGGCCAACTGCTGATCTGGCCAGCGCGGCAACC 60
DB 1 AAGGAAGTGGAGAACGAGAGAGCTGGCCAACTGCTGATCTGGCCAGCGCGGCAACC 60
QY 61 ATGCGCGGCGCTGGCGCCAGCGCGCCCAAGCGCCCACTTACAGAGCTGGCCAGGTTGCG 120
DB 61 ATGCGCGGCGCTGGCGCCAGCGCGCCCAAGCGCCCACTTACAGAGCTGGCCAGGTTGCG 120
QY 121 GTGCAAGAGCTGATTTGCGCGCGTGGAGCTGGCCAGCTTGGCCATTTGCGCGCGAG 180
DB 121 GTGCAAGAGCTGATTTGCGCGCGTGGAGCTGGCCAGCTTGGCCATTTGCGCGCGAG 180
QY 181 CAGGTGATGATGATGCTCCGAAAGCATCAACAGAGAGCTGCTCAAGCTGGCAAGC 240
DB 181 CAGGTGATGATGATGCTCCGAAAGCATCAACAGAGAGCTGCTCAAGCTGGCAAGC 240
QY 241 AGCGTGGCCGAGCTGGCCAGAGCAATGACGTGATGATGATGATGATGATGATGATG 300
DB 241 AGCGTGGCCGAGCTGGCCAGAGCAATGACGTGATGATGATGATGATGATGATGATG 300
```

```

QY 301 GACACCTGGAGAAACCGCTACTTTTGAACCTGTGAAAGACGACGACGCGATC 360
DB 301 GACACCTGGAGAAACCGCTACTTTTGAACCTGTGAAAGACGACGACGCGATC 360
QY 361 GTCTGTGTGTGTTTCCATGCGCCCGCCGACCGGCACTGCGCGACGCACTGCTCACTG 420
DB 361 GTCTGTGTGTGTTTCCATGCGCCCGCCGACCGGCACTGCGCGACGCACTGCTCACTG 420
QY 421 TACAAACGCGTGTGCGTGTGCGCAAGAACTGCGCGCGCAAGGGCGTGTGTGAC 480
DB 421 TACAAACGCGTGTGCGTGTGCGCAAGAACTGCGCGCGCAAGGGCGTGTGTGAC 480
QY 481 ATGAACGACGAGATCCAGTCCGCGCGTGAAGTGAAGATCAATCAAGACCGAA 540
DB 481 ATGAACGACGAGATCCAGTCCGCGCGTGAAGTGAAGATCAATCAAGACCGAA 540
QY 541 GCCTTCAAGAGCGCTGTGGGCGCGCTGTGGCATGTGTGTGAAGGCAAGTCTACTGTTT 600
DB 541 GCCTTCAAGAGCGCTGTGGGCGCGCTGTGGCATGTGTGTGAAGGCAAGTCTACTGTTT 600
QY 601 CGCTGTGCGCGCAAGCGGCAACAGTCACTCCAGATTGACATCAAGCAATCAAGCAGC 660
DB 601 CGCTGTGCGCGCAAGCGGCAACAGTCACTCCAGATTGACATCAAGCAATCAAGCAGC 660
QY 661 CTGCGCCGAGTGAACATGCGCTTACAGTATGGAAGTCAAGCGGCGCTTCAAGAGCC 720
DB 661 CTGCGCCGAGTGAACATGCGCTTACAGTATGGAAGTCAAGCGGCGCTTCAAGAGCC 720
QY 721 CTGCGCAACAGACGCGCGCAAGGCGCTGATCAATGCGCGCAACGCGCAATGCTGTGTG 780
DB 721 CTGCGCAACAGACGCGCGCAAGGCGCTGATCAATGCGCGCAACGCGCAATGCTGTGTG 780
QY 781 TCGCGGTGTGTGCGAGCGCTGCAAGAGTGTGCGCAAGAGCGGCGTGAATCAATTCGTTG 840
DB 781 TCGCGGTGTGTGCGAGCGCTGCAAGAGTGTGCGCAAGAGCGGCGTGAATCAATTCGTTG 840
QY 841 TCAC---GTCAACAGGCGCGTTTCTGCGTGTGCGTGAAGCGCGAGCGCGCAACGAAAC 897
DB 841 TCAC---GTCAACAGGCGCGTTTCTGCGTGTGCGTGAAGCGCGAGCGCGCAACGAAAC 897
QY 898 GACTGGGTGTGTGCGCGCAACGCTGAACCGCGCAAGAGCGCGCATCTGTGCGATGTGTGCA 957
DB 898 GACTGGGTGTGTGCGCGCAACGCTGAACCGCGCAAGAGCGCGCATCTGTGCGATGTGTGCA 957
QY 901 GACTGGGTGTGTGCGCGCAACGCTGAACCGCGCAAGAGCGCGCATCTGTGCGATGTGTGCA 960
DB 901 GACTGGGTGTGTGCGCGCAACGCTGAACCGCGCAAGAGCGCGCATCTGTGCGATGTGTGCA 960
QY 958 ATGACCAAGACCCAGGACAGCAAGAGGCTGCAAGCGCATTTTCTGGGAATATCTGATTA 1014
DB 958 ATGACCAAGACCCAGGACAGCAAGAGGCTGCAAGCGCATTTTCTGGGAATATCTGATTA 1014

```

RESULT 3

```

US-09-252-991A-5793
/ Sequence 5793, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 5793
/ LENGTH: 1182
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5793

```

Query Match 73.2%, Score 742, DB 4, Length 1182;

Best Local Similarity 83.9%; Pred. No. 5,3e-151;
Matches 851; Conservative 0; Mismatches 160; Indels 3; Gaps 1;

```

QY 1 AAGGAAGTGAAGAACACGAGACGAGACGCTGTGTGATCTCTGCGCACCGCGCGAC 60
DB 1 AAGGAAGTGAAGAACACGAGACGAGACGCTGTGTGATCTCTGCGCACCGCGCGAC 60
QY 169 AAGGAAGTGAAGAACACGAGACGAGACGCTGTGTGATCTCTGCGCACCGCGCGAC 228
DB 169 AAGGAAGTGAAGAACACGAGACGAGACGCTGTGTGATCTCTGCGCACCGCGCGAC 228
QY 229 ATGCGCGCGCGTGTGCGCGCGCGCGCGCAACAGCGCACTTACAGGCTGTGCG 288
DB 229 ATGCGCGCGCGTGTGCGCGCGCGCGCGCGCAACAGCGCACTTACAGGCTGTGCG 288
QY 121 GTGCAACAGCTGATTTGCGCGCGTGTGCGAGTGTGCGCGCAATGTGCGCGAG 180
DB 121 GTGCAACAGCTGATTTGCGCGCGTGTGCGAGTGTGCGCGCAATGTGCGCGAG 180
QY 289 GTGCAACAGCTGATTTGCGCGCGTGTGCGAGTGTGCGCGCAATGTGCGCGAG 348
DB 289 GTGCAACAGCTGATTTGCGCGCGTGTGCGAGTGTGCGCGCAATGTGCGCGAG 348
QY 181 CAGGTGATGACATGCTCTCCGAAAGCATCAACAGACCTGTCTCAAGCTGTGAC 240
DB 181 CAGGTGATGACATGCTCTCCGAAAGCATCAACAGACCTGTCTCAAGCTGTGAC 240
QY 349 CAGGTGATGACATGCTCTCCGAAAGCATCAACAGACCTGTCTCAAGCTGTGAC 408
DB 349 CAGGTGATGACATGCTCTCCGAAAGCATCAACAGACCTGTCTCAAGCTGTGAC 408
QY 241 AGCGTGTGCGAGCTGTGCGCGCGCGCGCAATGAGTGTGATCACTGCGCGAC 300
DB 241 AGCGTGTGCGAGCTGTGCGCGCGCGCGCGCAATGAGTGTGATCACTGCGCGAC 300
QY 409 ACCGTGTGCGAGCTGTGCGCGCGCGCGCGCAATGAGTGTGATCACTGCGCGAC 468
DB 409 ACCGTGTGCGAGCTGTGCGCGCGCGCGCGCGCAATGAGTGTGATCACTGCGCGAC 468
QY 301 GACACCTGGAGAAACCGCTACTTTTGAACCTGTGAAAGACGACGACGCGATC 360
DB 301 GACACCTGGAGAAACCGCTACTTTTGAACCTGTGAAAGACGACGACGCGATC 360
QY 469 GACACCTGGAGAAACCGCTACTTTTGAACCTGTGAAAGACGACGACGCGATC 528
DB 469 GACACCTGGAGAAACCGCTACTTTTGAACCTGTGAAAGACGACGACGCGATC 528
QY 361 GTCTGTGTGTGTTTCCATGCGCCCGCCGACCGGCACTGCGCGACGCACTGCTCACTG 420
DB 361 GTCTGTGTGTGTTTCCATGCGCCCGCCGACCGGCACTGCGCGACGCACTGCTCACTG 420
QY 529 GTGTGTGTGTGTTTCCATGCGCCCGCCGACCGGCACTGCGCGACGCACTGCTCACTG 588
DB 529 GTGTGTGTGTGTTTCCATGCGCCCGCCGACCGGCACTGCGCGACGCACTGCTCACTG 588
QY 421 TACAAACGCGTGTGCGTGTGCGCGCAAGGCGCTGATCAATGCTGTGTG 480
DB 421 TACAAACGCGTGTGCGTGTGCGCGCAAGGCGCTGATCAATGCTGTGTG 480
QY 589 TACAAACGCGTGTGCGTGTGCGCGCAAGGCGCTGATCAATGCTGTGTG 648
DB 589 TACAAACGCGTGTGCGTGTGCGCGCAAGGCGCTGATCAATGCTGTGTG 648
QY 481 ATGAACGACGAGATCCAGTCCGCGCGTGAAGTGAAGATCAATCAAGACCGAA 540
DB 481 ATGAACGACGAGATCCAGTCCGCGCGTGAAGTGAAGATCAATCAAGACCGAA 540
QY 649 ATGAACGACGAGATCCAGTCCGCGCGTGAAGTGAAGATCAATCAAGACCGAA 708
DB 649 ATGAACGACGAGATCCAGTCCGCGCGTGAAGTGAAGATCAATCAAGACCGAA 708
QY 541 GCCTTCAAGAGCGCTGTGGGCGCGCTGTGGCATGTGTGTGAAGGCAAGTCTACTGTTT 600
DB 541 GCCTTCAAGAGCGCTGTGGGCGCGCTGTGGCATGTGTGTGAAGGCAAGTCTACTGTTT 600
QY 709 GCGTTCAAGAGCGCTGTGGGCGCGCTGTGGCATGTGTGTGAAGGCAAGTCTACTGTTT 768
DB 709 GCGTTCAAGAGCGCTGTGGGCGCGCTGTGGCATGTGTGTGAAGGCAAGTCTACTGTTT 768
QY 601 CGCGTGTGCGCGCAAGCGCGCAACGCTGATCAATGCGCGCAATGCTGTGTG 660
DB 601 CGCGTGTGCGCGCAAGCGCGCAACGCTGATCAATGCGCGCAATGCTGTGTG 660
QY 769 CGCGCAACGCTGATCAATGCGCGCAACGCTGATCAATGCGCGCAATGCTGTGTG 828
DB 769 CGCGCAACGCTGATCAATGCGCGCAACGCTGATCAATGCGCGCAATGCTGTGTG 828
QY 661 CTGCGCCGAGTGAACATGCGCTTACAGTATGGAAGTGTGGAAGGCAAGTCTACTGTTT 720
DB 661 CTGCGCCGAGTGAACATGCGCTTACAGTATGGAAGTGTGGAAGGCAAGTCTACTGTTT 720
QY 829 CTGCGCCGAGTGAACATGCGCTTACAGTATGGAAGTGTGGAAGGCAAGTCTACTGTTT 888
DB 829 CTGCGCCGAGTGAACATGCGCTTACAGTATGGAAGTGTGGAAGGCAAGTCTACTGTTT 888
QY 721 CTGCGCAACAGACGCGCGCAAGGCGCTGATCAATGCGCGCAATGCTGTGTG 780
DB 721 CTGCGCAACAGACGCGCGCAAGGCGCTGATCAATGCGCGCAATGCTGTGTG 780
QY 889 CTGCGCCGAGTGAACATGCGCTTACAGTATGGAAGTGTGGAAGGCAAGTCTACTGTTT 948
DB 889 CTGCGCCGAGTGAACATGCGCTTACAGTATGGAAGTGTGGAAGGCAAGTCTACTGTTT 948
QY 781 TCGCGGTGTGTGCGAGCGCTGCAAGAGTGTGCGCAAGAGCGGCGTGAATCAATTCGTTT 840
DB 781 TCGCGGTGTGTGCGAGCGCTGCAAGAGTGTGCGCAAGAGCGGCGTGAATCAATTCGTTT 840
QY 841 TCAC---GTCAACAGGCGCGTTTCTGCGTGTGCGTGAAGCGCGAGCGCGCAACGAAAC 897
DB 841 TCAC---GTCAACAGGCGCGTTTCTGCGTGTGCGTGAAGCGCGAGCGCGCAACGAAAC 897
QY 949 GCGCGGTGTGTGCGAGCGCTGCAAGAGTGTGCGCAAGAGCGGCGTGAATCAATTCGTTT 1008
DB 949 GCGCGGTGTGTGCGAGCGCTGCAAGAGTGTGCGCAAGAGCGGCGTGAATCAATTCGTTT 1008
QY 1009 TCGCAACGAGTGTGCGCGCGCGCGCGCGCGCAACAGCGCACTTACAGGCTGTGCG 1068
DB 1009 TCGCAACGAGTGTGCGCGCGCGCGCGCGCGCGCGCAACAGCGCACTTACAGGCTGTGCG 1068
QY 898 GACTGGGTGTGTGCGCGCAACGCTGAACCGCGCAAGAGCGCGCATCTGTGCGATGTGTGCA 957
DB 898 GACTGGGTGTGTGCGCGCAACGCTGAACCGCGCAAGAGCGCGCATCTGTGCGATGTGTGCA 957
QY 1069 GACTGGGTGTGTGCGCGCAACGCTGAACCGCGCAAGAGCGCGCATCTGTGCGATGTGTGCA 1128
DB 1069 GACTGGGTGTGTGCGCGCAACGCTGAACCGCGCAAGAGCGCGCATCTGTGCGATGTGTGCA 1128
QY 958 ATGACCAAGACCCAGGACAGCAAGAGGCTGCAAGCGCATTTTCTGGGAATATCTGATTA 1011
DB 958 ATGACCAAGACCCAGGACAGCAAGAGGCTGCAAGCGCATTTTCTGGGAATATCTGATTA 1011
QY 1129 ATGACCAAGACCCAGGACAGCAAGAGGCTGCAAGCGCATTTTCTGGGAATATCTGATTA 1182
DB 1129 ATGACCAAGACCCAGGACAGCAAGAGGCTGCAAGCGCATTTTCTGGGAATATCTGATTA 1182

```

RESULT 4

US-09-252-991A-5877/C
 ; Sequence 5877, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 5877
 ; LENGTH: 1335
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-5877

Query Match 73.2%; Score 742; DB 4; Length 1335;
 Best Local Similarity 83.9%; Pred. No. 5,4e-151;

Matches 851; Conservative 0; Mismatches 160; Indels 3; Gaps 1;

QY 1 AAGGAAGTGAAGAACGAGCAAGCTGCGCAAGTGTGATCTGCGCAACCGCGCGCAC 60
 DB 1239 AAGGAAGTGAAGAACGAGCAAGCTGCGCAAGTGTGATCTGCGCAACCGCGCGCAC 1180
 QY 61 ATGCGCGGCGCTGGCG 120
 DB 1179 ATGCGCGGCG 1120
 QY 121 GTGCGCAAGCTGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 DB 1119 GTGCGCAAGCTGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1060
 QY 181 CAGGTGATGCAATGCGCTCCGAAAGCATCACCAAGCAAGCAAGCTGCTCAAGTGGCAAGC 240
 DB 1059 CAGGTGATGCAATGCGCTCCGAAAGCATCACCAAGCAAGCAAGCTGCTGGAAGTGGCAAG 1000
 QY 241 ACGGTGCGCGAGCTGCG 300
 DB 999 ACGGTGCGCGAGCTGCG 940
 QY 301 GACACCTGGAAGAAACCGCTACTTTTGAACCTGTGAAAGACCGCAAGCGGATC 360
 DB 939 GACACCTGGAAGAAACCGCTACTTTTGAACCTGTGAAAGACCGCAAGCGGATC 880
 QY 361 GTGCGTGTGCTTCCATGCG 420
 DB 879 GTGCGTGTGCTTCCATGCG 820
 QY 421 TACCAAGCGGTGCG 480
 DB 819 TACCAAGCGGTGCG 760
 QY 481 ATGAACGACGAGTCCAGTCCCGCGCGTGAAGAGCAATGATCAATCAAGACCGAA 540
 DB 759 ATGAACGACGAGTCCAGTCCCGCGCGTGAAGAGCAATGATCAATCAAGACCGAA 700
 QY 541 GCGTTCAAGAGCGCTGGGCGCGCTGGGCGATGTGTGGAAGCGCAAGTGTGATGTTTC 600
 DB 699 GCGTTCAAGAGCGCTGGGCGCGCTGGGCGATGTGTGGAAGCGCAAGTGTGATGTTTC 640
 QY 601 GCGCTGCG 660
 DB 639 GCGCGACCGGTGAGGCG 580
 QY 661 CTGCGCCGAGTGAAGTGTGATGCTGATGGAAGTGTGATGGAAGTGTGATGGAAGTGT 720
 DB 579 CTGCGCCGAGTGAAGTGTGATGCTGATGGAAGTGTGATGGAAGTGTGATGGAAGTGT 520

QY 721 CTGCGCAAGAACG 780
 DB 519 CTGCG 460
 QY 781 TCGCGCGGTGCG 840
 DB 459 GCGCGCGGTGCG 400
 QY 841 TCAC---GTCAACGAGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 897
 DB 399 TCGCACGTGATGCG 340
 QY 898 GACTGGGTGCG 957
 DB 339 GACTGGGTGCG 280
 QY 958 ATGACCAAGACCGCGAGAGAGAGTGAAGCGCATTTCTGGAATATGTA 1011
 DB 279 ATGACCAAGACCGCGAGAGAGTGAAGCGCATTTCTGGAATATGTA 226

RESULT 5

US-09-252-991A-5773
 ; Sequence 5773, Application US/09252991A
 ; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 5773

; LENGTH: 2748

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-5773

Query Match 55.4%; Score 562.2; DB 4; Length 2748;
 Best Local Similarity 85.2%; Pred. No. 4e-112;

Matches 640; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

QY 264 CAATGACGTGATGCGATGCTGATCAATGCGCAACCGCAACCGCTGGAAGAAACCGCCTA 323
 DB 1 CGACGACGTGACGCGCATGTGATCAACCAAGTACCGCAACCGCTGGAAGAAACCGCCTA 60
 QY 324 CTTTTTGAACCTGCGGAAAGAACCGCAACCGCATCGTGTGCGTTCATGCGCGCC 383
 DB 61 CTTCTGACCTGCGTGTGCGGACACCGCAAGACCTATCGTGTGCGTGTGCGTGTGCGCGCC 120
 QY 384 CGGACCGCATGTGCG 443
 DB 121 GGGACCGCATGTGCG 180
 QY 444 CAACAGACCTGCG 503
 DB 181 CGACAGGTGCGCACCG 240
 QY 504 GCGTGAAGTGAAGTGCATCAATCAAGACCGAAGCGCTTCAAGAGCGCGTGGGCGCC 563
 DB 241 CGGCGACCGAGCAAGATGTGATCAATCAAGACCGAAGCGCTTCAAGAGCGCGTGGGCGCC 300
 QY 564 GCTGGGCGATGTGTGAAGGCAAGTGTGATGCTGCTGCGCGCGCGCGCGCGCGCGCGCG 623
 DB 301 GCTGGGCGATGTGTGAAGGCAAGTGTGATGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 624 GGTCAACTCGGAGTTCATCAAGCAAGTCAAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCTA 683

Db	361	GGTGAACCTCCGAACTTGCATCAACAGAGATCTCCGGCTGGCTCCGGTGGAAATCGCTA	420
Qy	684	CACCTATGGCAAGTCATCCGACACCGGCTTACAAAGGCTCTGGACACAAACGGCGCCAAAGC	743
Db	421	CACCTACGGGCAAGTCAAGGACAACCGCTTACAAAGGCTCTGGCCCAAGCGCGCGCCAAAGGC	480
Qy	744	GCTGATTCATGCGCGGACCGGCAATGGCTCGGTTGTCGTCCGAGGAGTGGTGCAGGCTTGCA	803
Db	481	GATCATTCATGCGCGGACCGGCAACGGCTTCGGTTGCCCGGCGCGTGGTGGCCAGCCCTTGCA	540
Qy	804	GGAGCTGCGGCAAGACGGCGTGCAGATCAATTGTTCTGCAC---GTCAACAGGCGGGTTT	860
Db	541	GGAACTGGGAGACAAAGGGGTGAGATCAATCGGCTCTTCGACACGTCAATGCGGGGGGCTT	600
Qy	861	CGTGCTGCGTAAACGCCGAGACGCCGACGACAAAGAACATGAGTCTGTGCCCAACGACTT	920
Db	601	CGTCTGCGCAGACGCCGACGACGCCGAGACAAAGAACATGAGATCTGTGCCCAACGACTT	660
Qy	921	GAAACCCGCAAGAAAGCCCCGCAATCTTGGCCATGCTGGCAATGACACAAAGACCCACAGACAGCA	980
Db	661	GAAACCCGCAAGAAAGCCGCAATCTTGGCCGCGGCTGCAGATGACCAAGACCCACAGACAGCAA	720
Qy	981	GGAGCTGCAAGCGCATTTTCTGGGAATACTGA	1011
Db	721	GGAGCTGCAAGCGAATCTTCTGGGAATACTGCA	751

RESULT 6

```

US-09-328-352-2336
/ Sequence 2336, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 2336
/ LENGTH: 1125
/ TYPE: DNA
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-2336

```

Query Match	Score	DB	Length
26.9%	273	4	1125

Best Local Similarity 56.24; Pred. No. 8.4e-50;
Matches 557; Conservative 0; Mismatches 425; Indels 9; Gaps 2

Qy	30	CAACGAGTGATACCTGAGCCACCGGCGGACCAATGCGCGGCGTGGCGGCAAGCGGGCAA	89
Db	135	CAATGTTGTTGTGTGCTACTGACGGATCACTTGTGTGTGGCGGCAAGCTCAGCMAA	194
Qy	90	CAGCGCACCTACAGGCTGCGCAAGGTTGGCTGCACAAGCTGATTTGCCGGCGTCCGGA	149
Db	195	TATGTGCAACTTAACGCGCGCAAAAGTTCCAGTTGATCTTAAATCATGCAAGTTCTTCA	254
Qy	150	GCTGCGCACTCTGGCCAAATGTGCGCGGCGAGGAGTAGACAGATCGCTCCGAAACAT	209
Db	255	AATTCAAGATTGGCGAATGTATCTGTATCCAGACATTACAAGTAGTTCTGAAAGTAT	314
Qy	210	CACCAACGACGACCTGCTCAAGCTGGCAAGCAGCTGGCCGAGCTGCCGACAGCAATGA	265
Db	315	TACTGACAAAGAAATTTTACAAATTTGCTGTCAATTAATGAATCTGTAAAAAGCCAAAC	374
Qy	270	CGTCGATGGCATGTGATCAACCATGGACGACACCTTGAAGAAACCGGCTACTTTT	329
Db	375	TGTAAATGGGGTGTGATTAACAAGGTACAAATCTTTAAGAAACAGCATTTTCTT	434
Qy	330	GAACCTCTGTGAAAAGACGACAAAGCCATGCTGTGCTCGTTTCATGCGCCCCGGCAC	388
Db	435	AAATCTGTGTTCATACGTATTAACCAATCGTACTGTGTGGCTCAATGCGCCCATCAAC	494

QY	CGCCATGTCGCGCCACCGGCATGCTCAACCTGTACCAAGCGGTGGCCGTGGCAGCAACAA	449
Db	495 TGCCTTTCAAGCAGATGTCACCTTAACTTTAATAGTCGTGTGCAATTGGCCGCTTCTGA	554
QY	450 GGACTGCGCGGCGGAGGCGCTGTGTGTGACCATGAACGACAGATCCAGTCCGGGCGTGA	509
Db	555 TGAATGCAGAAAAATTAAGGCGTTATGTGTTCTCATGAACGACTCTATTTTGTGTCGTGGA	614
QY	510 CGTGAAGCAATGCCATCAACATCAAGACCGAAGCTTCAAGAGCCGCTGGGCGCGCTGGG	569
Db	615 TGTAACTTAAGGCATTTAATCAATCAAAATGCTTTGTAAAGCCAAATGGGCTGTTTGGG	674
QY	570 CATGCTGTGTGAAGGCAAGTCGTACTGTGTTCCGCTGCCGCGCAAGCGCCACACGATCA	629
Db	675 TAACTTGTGTAAGGCAAAACCATTTGTGTTTAAAGCAATCTGTTAAAGTCAATACAAATGC	734
QY	630 CTCGAGTTGCACATCAAGCAGATCAGACG-----CTGCCCAAGTGCATGCGCTTA	683
Db	735 TTCAAAATTTAATATTTGAAATATTTAAAGGATGCTCTTCCGACAGATACAAATCGTTTA	794
QY	684 CAGCTAATGGCAAGTCAACGACACGGGCTTCAAGGCGCTGGCACAAGAACGGGCGCAAGGC	743
Db	795 TGTGTTCTGAATCTATGCTTCTGTATGCTTTATGAAGCAATATGCAAAAGCTGGCGCTTAAGC	854
QY	744 GCTGATCCATGTCGCGCACCGGCATATGCTCGTGTGCTCGCGGGTGTGTCAGGCTTGA	803
Db	855 GATTAATTCATGACAGGTACAGGTAAATGTTCTGTAGCGAAATATATGTTCCAACTTAA	914
QY	804 GGAAGTCG---GCAAGAACCGCGTCGACATCTCTGTCGTCACTGACATCAAGGCGGTTT	860
Db	915 AAACTTCATGACAAAAACGGTATTCAAATTAATTCGCTCATCGGCGCTTCCAAAGGTTT	974
QY	861 CGTGTCTCGTAAACGCGCAGACGCGCCAGACGAACAAGACGACTGGTCTGTGGCCACGACT	920
Db	975 TGTATTACGTGATGCAAGAACACTGACTTAATATATGTTGGGTAGCAGCTATGATTT	1034
QY	921 GAACCGCAGAAAGGCGCCGATCTCTGCGCATGTGTGGCAATGACCAAGAACCAAGACAGCA	980
Db	1035 GAATTCCTCAAAAAGCGGTCTTTCTTGTGGCGGTTTGCACTTACAAAACCAACGATGCGAA	1094
QY	981 GGAAGTCGAGCGCATTTTCTGGGAATACTGA 1011	
Db	1095 AGAATATCAAGCATGTTCTGGCAGTACTTA 1125	

RESULT 7

US-08-447-422-15
; Sequence 15, Application US/08447422

GENERAL INFORMATION:

APPLICANT: SHAMI, Ezekiel Y.
; DOMICILIARY: ADDIS ABABA

APPLICANT: RAM,TEESTINGH Mohabir

TITLE OF INVENTION: Use of Antibody/Antigen Interactions To

1 TITLE OF INVENTION: Protect or modulate Biological Activity

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS: Polay & Tardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington, D.C.

; COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; ***** INFORMATION *****
;

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: IIS/08/447 4232

FILED DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/081,410

FILING DATE: 22-JUN-1993
 APPLICATION NUMBER: US 07/938,505
 FILING DATE: 31-AUG-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/205,748
 FILING DATE: 21-JUN-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 17923/102 HYLL
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5359
 TELE: 904136
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1948 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IS-08-447-422-15

433 GCGTGGCCAGCAAGAGATCGCGCGGCAAGGCGTGTGATCATGACGACGAG 492
 505 AGCGTCGCGCGATCTCTATGCCGCGGAGACCGCGGATATGCTGCTTAACGACCGC 564
 493 ATCCAGTCCGCGCGTGAAGTGAAGCAATGATCAATCAAGACCGAACCTTCAAGAC 552
 565 ATCGGCGCGCGCGCTTGTGACCAAAATTAATGCACTTCGTGATACCTTCCGGGG 624
 553 GCGTGGG--GCGCGTGGGCAATGCTGCTGGAAGCAATGCTGCTGCTTCCGCTGCCG 609
 625 CCGGAGAGGCGTATCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 684
 610 GCCAGCGCGCAAGCGTCAATCCGAGTTCGATCAAGCAAGTCAAGCAAGTCCGCGCG 669
 685 GATTAATTCACACGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 744
 670 GTGACATCGCTTACAGTATGAGCAATGACGACGACGCTTACAGAGCGCTTACAGAG 729
 745 GTGATGATTTATTTATGTTATACAGAGACGATCCGAGTATATGATGATGCGGACG 804
 730 AACGCGCGCAAGCGCTGATTCATGCGCGCAACCGGCAATGCTGCTGCTGCTGCTG 789
 805 CATCATGCGCGATATATCTATACGCGGACCGGCGGCGGCTGCTGCTGCTGCTGCT 864
 790 GTGCGAGCGCTGCAAGAGCTGCGCAAGAGCGGCTGCAAGTATGCTGCTGCTGCTG 849
 865 GCGCGCGGATTAATAAAGCGAGAGCGGCGATGCTGCTGCTGCTGCTGCTGCTGCTG 924
 850 CAGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909
 925 GCGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
 910 GCGCGAGCTGTAACCGCGCAAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
 970 GCGGCTGCTGTAACCGCGCAAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1029
 970 CAGGAGAGAGCTGCAAGCGCTTCTGGAATATCTGA 1011
 1030 AAGATCGCGAGCTGATTCAGCAATTTCTCATCTACTGA 1071

RESULT 9
 US-09-094-435-3
 Sequence 3, Application US/0904435
 Patent No. 6251388
 GENERAL INFORMATION:
 APPLICANT: Donald L. Durden
 TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES
 TITLE OF INVENTION: ASPARAGINASE IN THE TREATMENT OF HUMAN
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 MB
 MEDIUM TYPE: Storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: FASTSEQ for Windows 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/094,435
 FILING DATE: Filed Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/049,085
 FILING DATE: June 9, 1997

APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Weidberg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 234/274
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELETYPE: 67-3510
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1133 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-09-094-435-3
 Query Match 21.6%; Score 219; DB 3; Length 1133;
 Best Local Similarity 52.8%; Pred. No. 3.8e-38;
 Matches 521; Conservative 0; Mismatches 460; Indels 6; Gaps 2;

29 CCACGTGTGATCTCTGCGCACCGGCGGACCATGCGCGCTGCGCGCAGCGCGCA 88
 128 CCCAAGTGAATCTTACGCGACAGAGGACCATGCTGCTTCCGGGGAATCAAGCTCA 187
 89 ACAGCGCACCTTACAGAGCTGCGAGCTTGGCTGCAAGTGAATGCGCGCTGCGCG 148
 188 AGAG---TAGTACTCTGCTGAGAGCTCACCGTGTGAATAGCTTGTGAGCGCTG 244
 149 AGCTGGCGCACTTGGCGCAATGCGCGGAGCAGGTGATGAGATGCGCTTCCGAAGA 208
 245 CCATCAAGCACTTACGCGACATCAAGGTGAGAGATCTCAAGCATTTGCTTCCAGAGA 304
 209 TCACCAAGCAGCACTGCTTCAAGCTGCGCAAGAGCGGTGCGCGCTGCGCGCAAG 268
 305 TGACGGGTAAAGGTGCTTAAATAGCAAGCGGTGCAATGAGCTTCTGCGCAAAAG 364
 269 AGTGTATGATCTGCTATCAATCCATGCGACCGACCTTGAAGAAACCGCTACTTT 328
 365 AGACCGAAGCGGTGATCAATCCATGAGAACTGACCAATGAGAAACCGCTTCTTCC 424
 329 TGAACCTGCTGGAAGAAGCCGACAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTG 388
 425 TCACCTTACGCTGAAAGCCAAAGACCTGCTGCTTCTGAGCGCGCATGCTTCAAGCT 484
 389 CCGCATGTCGCGCGCGCATGCTCAACCTGTAACAAGCGCGCTGCGCGCGCAACA 448
 485 CTTCAATGATGCTGATGCGCGCATGAATCTTAAAGCGCGTGAATGAGGATCAACA 544
 449 AGGATTCGCGCGCGCAAGCGCTGCTGCTGCAATGAAGAGAGATTCAGTCCGGCG 508
 545 AAGCTTCACTAAACAAAGAGGTGATGATGAATGAATTCACGCGCGCAAG 604
 509 AGGTGAGAGATGATCAACATCAAGACCGAAGCTTCA---GAGCGCTGCGCGCG 565
 605 AAGCAACAGCTTCAACACCGACGATCAATGCTTCTTCCGCGCAACAGGTAAGA 664
 566 TGGGATGCTGCTGAGAGCAAGTGTATGCTGCTGCTGCTGCTGCGCGCAAGCGCA 625
 665 TCGGCAAGCTATATATGAGCAAGTGAAGTATTCATCAATCCGTTGACCTCAACC 724
 626 TCACTCGAGTTCATCAAGCAAGTCAAGAGCTGCGCGCGCGCGCGCGCGCTGCA 685
 725 TTGCAAGTGAAGTATATATGCAAAATGAAAGCTCCCAAGATCAATATCTTTTACG 784
 686 GCTATGCAAGCTGACAGACCGGCTTCAAGAGCGCTGCGCAACAAGCGCGCAAGCG 745
 785 CTCACCGCGATATATGATGTTTATGATGATGAGCGCTTCAAGCAAGAGCAAGA 844
 746 TGATCATGCGCGCAAGCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 805
 845 TCATCCATGCAAGGCAATGAGCAATGGAACCTTCTTCTTGAATCAAAATGCTTGA 904

QY 806 AGCTGCGAAGACGGCGTGCAGATCATTCGTCGTCACGTCACAGGGCGGTTTCGTC 865
 DB 905 AAGCAGCCAAATCAGCCGAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 964
 QY 866 TCGGTAAGCGCGGAGCGCCGAGCAGACAGACGACTGGGTGTCGCGCCGAGCTGAACC 925
 DB 965 CCCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1024
 QY 926 CGCAGAAAGCGCGGATTCGCGGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 985
 DB 1025 CTCAAAAAGCCAGAGTCTTCTTATGTTAGCCCTCAAAAATGATGATGATGATGATGAT 1084
 QY 986 TCGAGCGCATTTCTGGAATACTGAT 1012
 DB 1085 TCCAAAAGATCTTCTCCACCTATAT 1111

RESULT 10
 US-09-252-991A-5837/c
 ; Sequence 5837, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO: 5837
 ; LENGTH: 270
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-5837

Query Match 20.4%; Score 207.2; DB 4; Length 270;
 Best Local Similarity 85.8%; Pred. No. 9.7e-36;
 Matches 230; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
 QY 321 CTACTTTTGAACCTGTCGTAAGAAACGACGACGACGATCGTCGTCGTCGTCGTCGTCGTC 380
 DB 270 CTACTTCTGACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 211
 QY 381 CCCCAGCAGCGGATGTCGCGGACGCGATGCTCAACCTGTACACGCGGTCGTCGTCGTCGTC 440
 DB 210 CCGGAGCAGCGGATGTCGCGGACGCGATGCTCAACCTGTACACGCGGTCGTCGTCGTCGTC 151
 QY 441 CAGCAACAGACGTCGCGGAGGAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 500
 DB 150 CGGCAACAGTCGCGGACGCGGAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 91
 QY 501 CGGCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 560
 DB 90 CGGCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 31
 QY 561 CCGGCTGGGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 588
 DB 30 CCGGCTGGGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 3

RESULT 11
 US-09-557-884-1
 ; Sequence 1, Application US/09557884
 ; Patent No. 6506581
 ; GENERAL INFORMATION:
 ; APPLICANT: Fleischmann et al.
 ; TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments

; thereof, and Uses Thereof
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: MD
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3 1/2 inch diskette
 ; COMPUTER: Dell Pentium
 ; OPERATING SYSTEM: MS DOS v6.22
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/557,884
 ; FILING DATE: 25-Apr-2000
 ; CLASSIFICATION: <unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/476,102
 ; FILING DATE: JUN-5-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Michelle S. Marks
 ; REGISTRATION NUMBER: 41,971
 ; REFERENCE/DOCKET NUMBER: PB186P3
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 301-309-8504
 ; TELEFAX: 301-309-8439
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1830121 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-557-884-1

Query Match 13.6%; Score 138; DB 4; Length 1830121;
 Best Local Similarity 48.5%; Pred. No. 6.1e-20;
 Matches 474; Conservative 0; Mismatches 491; Indels 12; Gaps 3;
 QY 40 ATCTGCGACCGCGGACGACGATCGCGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 99
 DB 802735 ATCTGCGACCGCGGACGACGATCGCGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGG 802791
 QY 100 TACGAGCTGCGCAAGTTGCGTCGACGATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 159
 DB 802792 TATTAAGCTGACATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 802851
 QY 160 CTGGCAATGTCGCGGCGGACGAGTATGCGATGCGCTCCGAAAGCATGACCAAGCAGC 219
 DB 802852 ATTGCAACATTAAGGAGGACCAATGTTAATAAGTTTCAACAAGCATGATGATGACGA 802911
 QY 220 GACCTGCTCAAGCTGCGCAAGAGCGTGGCCGAGCTGCGCGGACGACGATGATGATGATGATG 279
 DB 802912 GTCTGGCTAAATCGGCAAAAGCCATCAATGCTCAATGTAATAAGTACT-----GATGGA 802965
 QY 280 ATGTCATCAACCATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 339
 DB 802966 TTGTGATTAACCATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 803025
 QY 340 GAAAGACGACCAAGCCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 399
 DB 803026 GTAAATGTGAAAACCGGTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 803085
 QY 400 GCCAGCGCATGCTCAACCTGTACAGCGCGTGGCGGTCGCGGACGACGACGACGACGACGACGAC 459
 DB 803086 GCTGATGCGCATTAATCTTACATGCTGCTGTCGCGAGCAGACGACGACGACGACGACGACGAC 803145
 QY 460 GCGAAGCGCGTGTGTGACATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 519
 DB 803146 GGTGCTGTGTTTATGTCGCAATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 803205

Qy	520	TCGATCAACATCAAGACCGAAGCCTTCAAGAGGC---	CGG66GCCCGCTGCGGCATGCGT	576
Db	803206	ACCAATACGACCCGACGTGCMAACGTTCCATTCAACCAAAATTATGTTCTCTAGGCTATATTT		8032855
Qy	577	GTGGAAGGCAAGTCGTACTGTTGTTCCGCGCTGCCGCAAGCGCCACACGCGTCAACTCCGAG		636
Db	803266	CATAACAGCAAGTGTGACTATGAAAGCTTCCCGAAGAAAGCAACATACATCAACACTCCG		8033255
Qy	637	TTGCACATCAAGCAGATCAGACGCTGCGCCGAGGTGACATCGCTTACAGCTATGCGCAAC		696
Db	803326	TTTAAACGTAGAAAATTAATAGATGAGCTTACCCAAAGTGGGATTAATATGCTTATTCAAAT		8033855
Qy	697	GTCAACGACACGCGCCTTACAGGCGCCCTGCGACAGAAAGCGGCCCAAGCGCTGATCCATGCC		756
Db	803386	GCACTCTGTGCAACCAATTAAACCATTACTCAATGCTGCGCTATCAAGGAGATTGATCTGCA		8034455
Qy	757	GGCACCGGCAATGGCTCGGCTGTCGTGCGGGGTGTCGACGCCCTGAGAGAGCTGCGCAAG		816
Db	803446	GGAGTTGGGAATGGAATATTTAATGTGCACACTTAGATGCGCTTAGAAAAAGCCGCAAA		8035055
Qy	817	AACGCGCTGCAGATCATTTGTTGTCATCAAGGCGCGTTTCGTGCTGCCTAAGCGC		876
Db	803506	GATAGCGTCTGTTTATGTCCTGTTCTTCTGCTGTACCAAGGGTTATTAACAATCTGTACGCT		8035555
Qy	877	GAGCAGCCCGACGACAGAAGACGACTGGGTCGTGGGCCACGACTGTAAACCGCAGAAAGGCT		936
Db	803566	GAAGTGTGATGATAGTAATAATATATGCGCTTTGTAGATCATCGTACTTTAATATCACAGAAAGCC		8036255
Qy	937	CGCATCTCGGCGCATGTGGCAATGACCAAGACCCAGACAGCAAGGAGCTGCAAGCGCAAT		996
Db	803626	CGCGGCTCTCTTCGCAATTAGCTTTAACTCAAACTAAGATCCNAAAGTAATTAATCAACATAT		8036855
Qy	997	TTCTGGGAATACTGATTA 1013		
Db	803686	TTCCGAAGACTTCTAAGA 803702		

RESULT 12
US-09-643-990A-1
Sequence 1

Sequence 1, Application US/09643990A
Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
Mark D. Adams
Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus Influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
/

Accession	Sequence	Position
Db	803506 GATAGCGTCGTGTAGTCCGTTCTCTCGTGTACCAACGGGTATATACAATCGTGAAGCT	803565
Qy	877 GAGCAGCCCGACGACACAGACGACTGGTCTGTGCCACGACTGAACCCGACAGAGGCC	936
Db	803566 GAACTGTGATGATAGTAAATATATGGCTTTGTACATCAGCTACTTTAAATCCACAGAAAAGCC	803625
Qy	937 CGCATTCCTGGGAGATGCTGGCAATGACCAAGACCCAGACACAGAGAGCTGCAGCGCATTT	996
Db	803626 CGCGTGTCTTGGCAATTAGCTTTTAACCTAAAGTATCCNAAGTAAATTCACCAATAT	803685
Qy	997 TTCTGGGAATCTGATA 1013	
Db	803686 TTCGAAGACTTCTAAGA 803702	

RESULT 13
US-09-543-681A-1219
; Sequence 1219, Application US/09543681A

```

1  APPLICANT: GARY BERTON
2  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
3  TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
4  FILE REFERENCE: 2709.1002-001
5  CURRENT APPLICATION NUMBER: US/09/543,681A
6  CURRENT FILING DATE: 2000-04-05
7  PRIOR APPLICATION NUMBER: US 60/128,706
8  PRIOR FILING DATE: 1999-04-09
9  NUMBER OF SEQ ID NOS: 8344
10 SEQ ID NO 1219
11 LENGTH: 1041
12 TYPE: DNA
13 ORGANISM: Proteus mirabilis
14 OS-09-543-681A-1219

```

Query Match	12.7%	Score 129;	DB 4;	Length 1041;
Best Local Similarity	47.2%;	Pred. No. 9.9e-19;		
Matches 464;	Conservative	0;	Mismatches 510;	Indels 9;
			Gaps	2

QY 29 CCAAGCTGCTATCTCTGGCCACCGGCGGACATCGCCGGCGCTGGCCGACCGCCGCA 88
 Db 68 CTAAATGTACAAATCCTCGCAACTGGGGGAAACATGCGAGGGGGTGTG---ACTGTGCA 124
 QY 89 ACAGGCGCACCTAACAGGCTGCGCAAGGTTGGCTGTGACAACTGATTTGCCGGCTGGCG 148
 Db 125 CTCTCTCAAGCTTACACCGCAGAAAAATTAGGCATTGATACATTGATTATCCGTACTGT 184
 QY 149 AGCTGGCCGACCTGGCCCAATGTGCGCGGCGAGCGAGTGAATGCAAGATGCGCTCCGAAAGCA 208
 Db 185 AGGCAAAAAAGTGTCTAACTTAACAGGCGAGCAAGTGTTAATATGTGGTTGCGCAAGTA 244
 QY 209 TCACCAACGACGACCTGTCTAAGCTGTGCAAGCTGACAGGCTGACCGACGACGACATG 268
 Db 245 TGAAGACCAAGTGTGTTAAACTGGCGAATAAAT-----TAATACGATTTGTGATA 298
 QY 269 ACGTGAATGGATGCTCATCAACCCATGGCACCGACACCTGTGAAAGAAACGGCTTACTTTT 328
 Db 299 AAACAGACGGCTTTGTGTGATTAACCCATGGTACAGATACGATGAAAGACCGCTATTTC 358
 QY 329 TGAACCTGTGAAAAAGACCGACAGCCGATGCTGTGTGTGGTTCCATGGCCCCCGCA 388
 Db 359 TTGACTTAACACAGCATGTAAAAAACGGGTGTGATGTATGGGGGCAATGGGCCACAGAA 418
 QY 389 CCGCATGTCCGCGGACGGCATGTCTCAACTGTACAAAGCCGTGGCGGTGGCCAGCAACA 448
 Db 419 CAGCACTTAAGTGTGATGTGTCCATTAACTCTTAATATGCCGTGTGGTGGCAAGTGATA 478
 QY 449 AGGACTGCGCGGCAAGGCGTGTCTGTGTGACCATGAACGACGAGATCAAGTCCGCGCGTG 508
 Db 479 AAGACTTGAAAAATGTGTGTGATTAAGTTACATGAATAAATTCAAGTATCAGTGTAAAG 538
 QY 509 ACGTAGCAAGTGCATCAACATCAAGACCGAAAGCTTCAAGAGCGGCTGGGGCCCGCTGG 568

Db	539	ATGTTGTCAAAATAATACACAGAAAGTCAGACATTCAGCCGCAATTAATGACAGTGTCT	598
Qy	569	GCATGATGATGGAAGAGCAAGTCGATCTGCTTCGCGCCGAGCGCCACACGGTCA	628
Db	599	AAGGTTATGTTTCATGATGTAAAGTACACTTATACACTGTGCAACACCGCGCTGAGA	658
Qy	629	ACTCCGAGTTTGCACATCAGACAGATTCAGCAGCCTGCCCAGGTGACATCGCTTACAGCT	688
Db	659	AAGTTGCTTTGATAGTACAGAAATTAACCGAATTACCAAAAGATGATATTGTTATTAAC	718
Qy	689	ATGCGAACATCACCAGACACGGGCTTACAAAGGCCCTGCGACACAGAAAGCGGGCCAAAGCGCTGA	748
Db	719	ATGCTAATGATCATGATTTTACCGGCTTAAAGCGTTTATAGATAACCACTTTAAAGTATGTG	778
Qy	749	TCCATGCGCGCACCGGCAATGAGTGCCTGGTGTCTGTCGCGGAGTGCACAGCCCTGCAGAGNC	808
Db	779	TTAGTGCAGGTGTGCGCAATGTAATCTTAATTTCTGACATCTTAAATATCTTTAGCTAGTG	838
Qy	809	TGCGCCAGAAACGGCGTGCAGATCATTTGTTGTCACGTCAACAGGCGGTTTCTGTCTGC	868
Db	839	GTTGTAATAAAGAGTGTGTGTGTGTTGTTTCAGTCTGTGTTCGCGTGGATTTTACACAC	898
Qy	869	GTAACGCCGACAGCCCCGACGACAAAGAACGCTGGTCTGTGCGCCACGACTGTAACCCGC	928
Db	899	AAAATGCTGAAGTATGATGATGCAAAAATATGGCTTTATTTGCTTCAGAAAGTTTAATCCAC	958
Qy	929	AGAAAGCCCGCATCTTGCGCATGTGTGCAATGACCAAGACCCAGACACAGCAAGAGCTGC	988
Db	959	AGAAAGCAAGAGTGTATTTCACAGTTATCTTTGACAGAAACAGATATCCGCGCACGATCC	1018
Qy	989	AGCGCATTTTCTGGGAATCTGCA	1011
Db	1019	AAGAAACTTTGAATAATATTTAA	1041

RESULT 14
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328

```

? GENERAL INFORMATION:
? APPLICANT: FLEISCHMAN, Robert D.
? APPLICANT: WHITE, Owen R.
? APPLICANT: PRASER, Claire M.
? APPLICANT: VENTER, John C.
? TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
? TITLE OF INVENTION: TUBERCULOSIS
? FILE REFERENCE: 24366-20007.00
? CURRENT APPLICATION NUMBER: US/09/103,840A
? CURRENT FILING DATE: 1998-06-24
? NUMBER OF SEQ. ID NOS.: 2
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 2
?
? LENGTH: 4403765
?
? TYPE: DNA
?
? ORGANISM: Mycobacterium tuberculosis
? FEATURE:
? OTHER INFORMATION: CDC 1551
? OTHER INFORMATION: "n" bases at various positions throughout the sequence
? OTHER INFORMATION: represent a, t, c or g
?
? US-09-103-840A-2

```

	Query Match	Best Local Similarity	8.2%; Score 83; DB 3; Length 4403765;
	Query	164; Conservative	0; Mismatches 135; Indels 0; Gaps 0;
	Db	1742101	230 AGCTGCGACAGCGCTGCGCGAGCTGGCCGACAGCAATGACGTGATGCGATCGTCATCA 289
	Qy	290	CCCATGGCAGCGACACCCCTGGAGAAACCGCTACTTTTGAACCTGCTGGAAAAAGACG 349
	Db	1742041	CTCAGCGCACCGAATACGCTCTGAAAGAGACCGCTCTGTGGCTGGACCTACCTATGCGGATA 1741982

QY	350	ACAGGCGGATGGTGGTGGATTCAATGGGCGCCGGACCGCAATGTCGGCGGACGGCA	409
Db	1741981	GCCGCGCGGTGTGTGTTGACCCGAGGCGCAATGCTCAAGCGCGCATGTCCCGCGCGCTAGTGC	1741922
QY	410	TGCTCAACTGTACAGACGCGTGGCCCGTGGCGACAGCAAGAGACTGGGCGCGGACGGGCG	469
Db	1741921	CGGCGCAATCTGCGGACGCGGTTTGCGGTGCGCGCGGACCCAGCGCGCCGCGACCTCGGGG	1741862
QY	470	TGCTGTGACCATGACAGACGAGATCCAGTCCGGGCGTGAAGTGAAGCAAGTCATCAAC	528
Db	1741861	TGCTGTGACGTTTTCGGCGGCGCGGGTGTTCAGACCTTTGGGCGCTGACAGATGGGCCAAC	1741803

RESULT 15

```

US-09-103-840A-1/c
/ Sequence 1, Application US/09103840A
/ Patent No. 6294328
/ GENERAL INFORMATION:
/ APPLICANT: FLUBSCHMAN, Robert D.
/ APPLICANT: WHITE, Owen R.
/ APPLICANT: FRASER, Claire M.
/ APPLICANT: VENTER, John C.
/ TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
/ TITLE OF INVENTION: TUBERCULOSIS
/ FILE REFERENCE: 24366-20007.00
/ CURRENT APPLICATION NUMBER: US/09/103.840A
/ CURRENT FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 4411529
/ TYPE: DNA
/ ORGANISM: Mycobacterium tuberculosis
/ OTHER INFORMATION: H37Rv
/ US-09-103-840A-1

```

Query Match	8.2%	Score 83	DB 3	Length 4411529
Best Local Similarity	5.8%	Pred. NO.	5.5e-08	
Matches 164	Conservative	0	Mismatches 135	Indels 0
				Gaps 0

QY	230	GGCTGGCAAGAGGGTGGCCGAGCTGGCCGACACATATGACGTGAAATGGCATGGTCATCA	289
Db	1741976	ACCGGATATGGCGCTGTCACTGACAGAGGATTTCCGGCGCGCGACGATGGGGTGGTCAATCA	1741917
QY	290	CCCATGGCACCAGACCCCTGGAGAAAACCGCTTACTTTTGAACCTTGTGGAAAAGACG	349
Db	1741916	CTCACGGCACAACGATACGCTTCAGAAAGACCGCTGTGTGGCTGAGACCTCACTATACGGGTA	1741857
QY	350	ACAAAGCGGATGTGTGTGTGGTTCCATATGCGCCCGGACACCGCATATCGCCGACAGGCA	409
Db	1741856	GCGCGCCGGTGTGTATTACCGGGGCGATGTCAAGCGCGCATGTCCCGCGCGCGATGACC	1741797
QY	410	TGCTCAACTGTACAAAGCGCGTGGCCGTGGCCAGCAACAAGACTCGCGCGACAGGGCG	469
Db	1741796	CGGCAATCTTGTGGGAGCGCGCTTGTGGGTGCGCGCGAACCCAGCGCGCGGAGACTGGGGG	1741737
QY	470	TGCTGTGTGACATGAAGAGCAAGATCCAGTCCGGGCGTGAAGCTGAGCAATTCATCAAC	528
Db	1741736	TGCTGTGTGATGTTTGGCGCGCGCGGGTGTATACGCTTTTGGGCGCTTGACACAAGGTGGCCAAC	1741678

Search completed: March 22, 2005, 12:57:48
Job time : 227 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 21:10:06 ; Search time 180 Seconds
(without alignments)
955.881 Million cell updates/sec

Title: US-09-842-628-2
Perfect score: 1680
Sequence: 1 KEVENQGLKLVAVLATGST.....MVAMTKTQDSKELQRIFMWEY 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1592.5	94.8	362	1	ASPO_PSEPK
2	1579.5	94.0	337	1	ASPO_PSEB7
3	1480.5	88.1	362	1	ASPO_PSEFL
4	1432.5	85.3	362	1	ASPO_PSEB8
5	1040.5	61.9	355	2	Q6PAL6
6	1037.5	61.8	331	1	ASPO_ACTGL
7	795	47.3	354	2	Q87J79
8	786	46.8	348	2	Q8XC02
9	786	46.8	348	2	Q83085
10	785	46.7	348	2	Q8CVR4
11	782	46.5	352	2	Q8A446
12	772	46.0	348	1	ASG2_BCOLI
13	768	45.7	352	2	Q64NH4
14	767	45.7	378	2	Q7VIR4
15	766	45.6	354	2	Q6L134
16	761	45.3	348	2	Q8XG73
17	761	45.3	348	2	Q7CPT7
18	758	45.1	349	1	ASG2_HABIN
19	758	45.1	375	1	ASG2_BACSU
20	757	45.1	330	1	ASPG_NOLSU
21	749.5	44.6	345	1	Q8ZGB7
22	749	44.6	348	1	ASPG_ERWCH
23	748.5	44.6	345	2	Q6GCJ2
24	739.5	44.0	331	2	Q9PJ82
25	735	43.8	346	2	Q6QAP4
26	735	43.8	349	2	Q7MVK3
27	734	43.7	348	2	Q6QAP3
28	730	43.4	349	2	Q7VNV6
29	729	43.4	355	2	Q65GV3
30	717	42.7	330	1	ASPG_HILPY
31	687	40.9	332	1	ASPG_HILPY

32	685.5	40.8	347	2	Q7N6D9	Q7N6D9
33	681	40.5	347	2	Q8ZLD3	Q8ZLD3
34	679	40.4	347	2	Q6ZJ01	Q6ZJ01
35	679	40.4	347	2	Q63U26	Q63U26
36	631	37.6	379	2	Q89KF7	Q89KF7
37	600	35.7	333	2	Q7WRB6	Q7WRB6
38	594	35.4	351	2	Q631L0	Q631L0
39	592	35.2	333	2	Q7W2H6	Q7W2H6
40	557	33.2	367	2	Q8XZT9	Q8XZT9
41	552.5	32.9	349	2	Q6CZM5	Q6CZM5
42	550	32.7	356	1	ASG2_SCHPO	Q9UB77
43	550	32.7	356	1	ASG4_SCHPO	Q8TIF8
44	527	31.4	360	1	ASG3_SCHPO	Q8NXC0
45	524	31.2	360	1	ASG1_SCHPO	P87015

ALIGNMENTS

RESULT 1
ASPO_PSEPK STANDARD; PRT; 362 AA.
ID ASPO_PSEPK
AC Q88K39;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Glutamine-asparaginase precursor (EC 3.5.1.38) (L-asparagine/L-glutamine amidohydrolase) (L-ASNase/L-GLNase).
GN Name=anab; OrderedLocustNames=PP2453;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160486;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.B., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H., Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M., Brinac L.M., Beaman M.J., Deboy R.T., Daugherty S.C., Kolony J.F., Madupu R., Nelson W.C., White O., Peterson J.D., Knout H.M., Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Ryan K., Moazzez A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D., Medler H., Lauber J., Stjepandic D., Hohenstein J., Stratz M., Heim S., Kiewitz C., Eisen J.A., Timmls K.N., Duesterhoef A., Tuenmler B., Frazer C.M.;
RT Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440.";
RL Brivion. Microbiol. 4:799-808(2002).
CC -1- CATALYTIC ACTIVITY: L-glutamine + H(2)O = L-glutamate + NH(3).
CC -1- CATALYTIC ACTIVITY: L-asparagine + H(2)O = L-aspartate + NH(3).
CC -1- SUBUNIT: Homotrimer (by similarity).
CC -1- SUBCELLULAR LOCATION: Periplasmic (by similarity).
CC -1- SIMILARITY: Belongs to the asparaginase 1 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC
EMBL, AS016783; AAN68065.1; -.
HSSP, P10182; 4PGA.
DR TIGR, PP2453; -.
DR InterPro, IPR004550; Asnase II.
DR InterPro, IPR006034; Asp/Glutamase.
DR Pfam, PF00710; Asparaginase; 1.
DR PRINTS, PR00139; ASNGNAB.
DR PRODOM, PD003221; Asp/Glutamase; 1.
DR PROSITE, PS00144; ASN_GLN_ASE_1; 1.
DR PROSITE, PS00917; ASN_GLN_ASE_2; 1.
KW Complete proteome; Hydrolase; Periplasmic; Signal.

FT SIGNAL 1 25 Potential.
 FT CHAIN 26 362 Glutamine-asparaginase.
 FT ACT SITE 45 45 By similarity.
 FT ACT SITE 125 125 By similarity.
 FT ACT SITE 126 126 By similarity.
 FT ACT SITE 198 198 By similarity.
 SQ SEQUENCE 362 AA, 38608 MW, C33F185D8Y1053F0 CRC64,
 Query Match 94.8%; Score 1592.5; DB 1, Length 362;
 Best Local Similarity 94.1%; Pred. No. 9, 2e-90;
 Matches 317, Conservative 9, Mismatches 10, Indels 1, Gaps 1,
 QY 1 KEVENOQKIANVYIATGTTIAGAGSAANSATYQAAKGVDELKAGVELADLANVGE 60
 DB 26 KEAETQOKIANVYIATGTTIAGAGSAANSATYQAAKGVDELKAGVELADIANVGE 85
 QY 61 QVMQIASSEITDDILKLASSVAELADSDVGIIVTHDTPLTEFRAVPLINVEKDKRI 120
 DB 86 QVMQIASSEISNDLKLKGRVAELAESKDVGDIVITHTDTLEAFPLINVEKDKPI 145
 QY 121 VVVGSRPCTAMSDGMLNLYNAVAASNDKSGKGLVTNDELQSGDVEKINIKTE 180
 DB 146 VVVGSRPCTAMSDGMLNLYNAVAASNDKSGKGLVTNDELQSGDVEKAVNIKTE 205
 QY 181 AFKSAMGPIGMVVEGKSYFRLPAKHTVNSEFDIKQISLPQVDIAYGYGVDTTAYYA 240
 DB 206 AFKSAMGPIGMVVEGKSYFRLPAKHTVNSEFDIKQISLPQVDIAYGYGVDTTAYYA 265
 QY 241 LAONGAKALIHAGTGSVSRVPLQELRNKGVQIITSSR-QQSGFVLRAAEQDDDN 299
 DB 266 LAONGAKALIHAGTGSVSRVPLQELRNKGVQIITSSRQSGFVLRAAEQDDDN 325
 QY 300 DWVAHDINPQKARILAMVAMTKTQDSKEIQRIFWEX 336
 DB 326 DWVAHDINPQKARILAMVAMTKTQDSKEIQRIFWEX 362
 RESULT 2
 ASFO_PSES7
 ID ASFO_PSES7 STANDARD, PRT, 337 AA.
 AC P10182;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DB Glutamine-asparaginase (EC 3.5.1.38) (L-asparagine/L-glutamine
 DB amidohydrolase) (L-ASNase/L-GLNase) (PGA).
 GN Name:anab;
 OS Pseudomonas sp. (strain 7A).
 OC Bacteria; Proteobacteria.
 OC NCBI_TaxID=313;
 OK NCBI_TaxID=313;
 RN [1]
 RP SEQUENCE, AND X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=94347759; PubMed=8068664;
 RA Lubkowski J., Wlodawer A., Ammon H.L., Copeland T.D., Swain A.L.;
 RT "Structural characterization of Pseudomonas 7A glutamine-
 RT asparaginase.";
 RT Biochemistry 33:10257-10265(1994).
 RN [2]
 RP SEQUENCE OF 1-26.
 RX MEDLINE=78080774; PubMed=619999;
 RA Holtenberg J.B., Ericsson L., Roberts J.;
 RT "Amino acid sequence of the diazooxonorleucine binding site of
 RT Acinetobacter and Pseudomonas 7A glutamine-asparaginase.";
 RT Biochemistry 17:411-417(1978).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
 RX MEDLINE=97172924; PubMed=9020792; DOI=10.1021/bi961979x;
 RA Jakob C.G., Lewinski K., Lacomte M.W., Roberts J., Leboda L.;
 RT "Ion binding induces closed conformation in Pseudomonas 7A
 RT glutamine-asparaginase (PGA): crystal structure of the PGA-SO4(2-)-
 RT NH4+ complex at 1.7-A resolution.";
 RT Biochemistry 36:923-931(1997).
 CC -1- CATALYTIC ACTIVITY: L-glutamine + H(2)O = L-glutamate + NH(3).
 CC -1- SUBUNIT: Homotrimer.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: Belongs to the asparaginase 1 family.
 DR PDB, 1D0I; X-ray; A/B=8-337.
 DR PDB, 1D0P; X-ray; A/B=8-337.
 DR PDB, 3PGA; X-ray; 1/2/3/4=1-337.
 DR PDB, 4PGA; X-ray; A/B=1-337.
 DR InterPro; IPR004550; Asnase_II.
 DR InterPro; IPR006034; Asp/Glutamase.
 DR Pfam; PF00710; Asparaginase; 1.
 DR PRINTS; PR00139; ASNGLNASE.
 DR ProDom; PD003221; Asp/Glutamase; 1.
 DR TIGRFAMs; TIGR00520; asnase_II; 1.
 DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
 DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
 KW 3d-structure; Direct protein sequencing; Hydrolase; Periplasmic.
 FT ACT_SITE 20 20
 FT ACT_SITE 100 100
 FT ACT_SITE 101 101
 FT ACT_SITE 173 173
 FT ACT_SITE 173 173
 FT CONFLICT 7 7
 FT CONFLICT 12 12
 FT CONFLICT 11 18
 FT HELIX 20 22
 FT STRAND 24 25
 FT TURN 28 29
 FT STRAND 34 35
 FT HELIX 41 46
 FT TURN 47 47
 FT TURN 49 50
 FT TURN 51 54
 FT HELIX 56 65
 FT STRAND 67 69
 FT HELIX 72 87
 FT TURN 89 90
 FT STRAND 93 97
 FT TURN 101 102
 FT HELIX 103 113
 FT STRAND 120 123
 FT TURN 129 130
 FT TURN 132 133
 FT HELIX 136 147
 FT TURN 148 148
 FT HELIX 150 152
 FT TURN 153 154
 FT STRAND 158 159
 FT TURN 162 163
 FT STRAND 164 167
 FT TURN 168 170
 FT STRAND 172 173
 FT STRAND 175 175
 FT TURN 179 180
 FT STRAND 182 183
 FT TURN 185 186
 FT TURN 190 193
 FT TURN 194 195
 FT STRAND 196 199
 FT STRAND 208 210
 FT HELIX 215 217
 FT STRAND 224 228
 FT TURN 231 232
 FT HELIX 236 243
 FT TURN 244 245
 FT STRAND 248 254
 FT TURN 255 257
 FT STRAND 258 258
 FT TURN 261 263
 FT HELIX 264 272
 FT TURN 273 274
 FT STRAND 276 281
 FT TURN 286 286
 FT TURN 291 293

O -> H (in Ref. 2).
 V -> R (in Ref. 2).

FT HELIX 297 300
 FT TURN 301 301
 FT STRAND 303 304
 FT TURN 306 307
 FT HELIX 310 320
 FT TURN 321 323
 FT HELIX 327 337
 SQ SEQUENCE 337 AA, 36200 MW, 11DEBCA67CB1A75B CRC64;

Query Match 94.0%; Score 1579.5; DB 1; Length 337;
 Best Local Similarity 94.1%; Pred. No. 5.3e-89;
 Matches 317; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

QY 1 KEVENOQKLANVYIATGTTAGAGASANSATYQAAYGVDELIAVPELDLANVRC 60
 DB 1 KEVENOQKLANVYIATGTTAGAGASANSATYQAAYGVDELIAVPELDLANVRC 60
 QY 61 QVMOJASESITNDLLKASSVAELADSDVDGIYTHGTDLLETFAYELAVETKDKPI 120
 DB 61 QVMOJASESITNDLLKASSVAELADSDVDGIYTHGTDLLETFAYELAVETKDKPI 120
 QY 121 VVYGSRRPTAMSDQMLLYNAVAASNKDSRGKGLVYTMDEIOSGRDVSINIKTE 180
 DB 121 VVYGSRRPTAMSDQMLLYNAVAASNKDSRGKGLVYTMDEIOSGRDVSINIKTE 180
 QY 181 AFKSAWGLGMYVEGKSYWFRLLPAKRHTVNSEFDIKQISLPQVDIAYSYGAVTDTAYKA 240
 DB 181 AFKSAWGLGMYVEGKSYWFRLLPAKRHTVNSEFDIKQISLPQVDIAYSYGAVTDTAYKA 240
 QY 241 LAQNGAKALIHAGTNGSVSSRVVPALQELRKNGVQIIRSSR-QQGGFYLRNAEQPDDKN 299
 DB 241 LAQNGAKALIHAGTNGSVSSRVVPALQELRKNGVQIIRSSR-QQGGFYLRNAEQPDDKN 299
 QY 300 DMVVAHDLNPKARKILAMVAMTKTODSKELQRIFWMEY 336
 DB 301 DMVVAHDLNPKARKILAMVAMTKTODSKELQRIFWMEY 337

RESULT 3

ASPO_PSEFL STANDARD; PRT; 362 AA.
 ID ASPO_PSEFL
 AC 068897;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Glutaminase-asparaginase precursor (EC 3.5.1.38) (L-asparagine/L-glutamine amidohydrolase) (L-ASNase/L-GLNase).
 GN Name=ansb;
 OS Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=294;
 RX NCB1_TaxID=294;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13525;
 RA Hueber A., Kleopener U., Roehm K.H.;
 RT "Cloning, sequence analysis, and gene expression of the P. fluorescens
 gene (ansb) encoding periplasmic glutaminase/asparaginase."
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: L-glutamine + H(2)O = L-glutamate + NH(3).
 CC -1- CATALYTIC ACTIVITY: L-asparagine + H(2)O = L-aspartate + NH(3).
 CC -1- SUBUNIT: Homotrimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: Belongs to the asparaginase 1 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

DR EMBL; AF056495; AAC33155.1; -.
 DR HSSP; P10182; 3PGA.
 DR InterPro; IPR004550; Asnase II.
 DR InterPro; IPR006034; Asp/Glutaminase.
 DR Pfam; PF00710; Asparaginase; 1.
 DR PRINTS; PR00139; ASNGINASE.
 DR ProDom; PD003221; Asp/Glutaminase; 1.
 DR TIGRFAMe; TIGR00520; asnase II; 1.
 DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
 DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
 KW Hydroxylase, Periplasmic, Signal.
 FT SIGNAL 1 25 Potential.
 FT CHAIN 26 362 Glutaminase-asparaginase.
 FT ACT_SITE 45 45 By similarity.
 FT ACT_SITE 125 125 By similarity.
 FT ACT_SITE 126 126 By similarity.
 FT ACT_SITE 198 198 By similarity.
 SQ SEQUENCE 362 AA, 38735 MW, E8C640D24C5FP31B CRC64;

Query Match 88.1%; Score 1480.5; DB 1; Length 362;
 Best Local Similarity 87.2%; Pred. No. 6.8e-83;
 Matches 294; Conservative 20; Mismatches 22; Indels 1; Gaps 1;

QY 1 KEVENOQKLANVYIATGTTAGAGASANSATYQAAYGVDELIAVPELDLANVRC 60
 DB 26 KEVENOQKLANVYIATGTTAGAGASANSATYQAAYGVDELIAVPELDLANVRC 85
 QY 61 QVMOJASESITNDLLKASSVAELADSDVDGIYTHGTDLLETFAYELAVETKDKPI 120
 DB 86 QVMOJASESITNDLLKASSVAELADSDVDGIYTHGTDLLETFAYELAVETKDKPI 145
 QY 121 VVYGSRRPTAMSDQMLLYNAVAASNKDSRGKGLVYTMDEIOSGRDVSINIKTE 180
 DB 146 VVYGSRRPTAMSDQMLLYNAVAASNKDSRGKGLVYTMDEIOSGRDVSINIKTE 205
 QY 181 AFKSAWGLGMYVEGKSYWFRLLPAKRHTVNSEFDIKQISLPQVDIAYSYGAVTDTAYKA 240
 DB 206 AFKSAWGLGMYVEGKSYWFRLLPAKRHTVNSEFDIKQISLPQVDIAYSYGAVTDTAYKA 265
 QY 241 LAQNGAKALIHAGTNGSVSSRVVPALQELRKNGVQIIRSSR-QQGGFYLRNAEQPDDKN 299
 DB 266 LAQNGAKALIHAGTNGSVSSRVVPALQELRKNGVQIIRSSR-QQGGFYLRNAEQPDDKN 325
 QY 300 DMVVAHDLNPKARKILAMVAMTKTODSKELQRIFWMEY 336
 DB 326 DMVVAHDLNPKARKILAMVAMTKTODSKELQRIFWMEY 362

RESULT 4

ASPO_PSEFL STANDARD; PRT; 362 AA.
 ID ASPO_PSEFL
 AC 091407;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Glutaminase-asparaginase precursor (EC 3.5.1.38) (L-asparagine/L-glutamine amidohydrolase) (L-ASNase/L-GLNase).
 GN Name=ansb; Ordered locus names=PA1337;
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=287;
 RX NCB1_TaxID=287;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RC MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
 RA Hickey C.K., Pham X.-Q.T., Eryin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brimham F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Coltry L., Tolentino B., Westbrock-Madman S., Yan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiser J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RA "Complete genome sequence of Pseudomonas aeruginosa PAO1, an

RT opportunistic pathogen." ;
 RL Nature 406:959-964(2000) .
 CC -1- CATALYTIC ACTIVITY: L-glutamine + H(2)O = L-glutamate + NH(3) .
 CC -1- CATALYTIC ACTIVITY: L-asparagine + H(2)O = L-aspartate + NH(3) .
 CC -1- SUBUNIT: Homotrimer (By similarity) .
 CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity) .
 CC -1- SIMILARITY: Belongs to the asparaginase 1 family .
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch) .
 CC -----
 CC EMBL, AB004563; ANG04726.1; --
 DR PIR, C83478; C83478.
 DR HSP, P10182; 3PGA.
 DR InterPro: IPR004550; Asnase II.
 DR InterPro: IPR006034; Asp/Glutamase.
 DR Pfam: PF00710; Asparaginase; 1.
 DR PRINTS; PR00139; ASNGLNASE.
 DR ProDom: PD003221; Asp/Glutamase; 1.
 DR TIGRFAMs; TIGR00520; asnase II; 1.
 DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
 DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
 DR Complete proteome: Hydratase; Periplasmic; Signal.
 KM Complete proteome: Hydratase; Periplasmic; Signal.
 FT SIGNAL 1 25 Potential.
 FT CHAIN 1 26 362 Glutaminase-asparaginase.
 FT ACT_SITE 45 45 By similarity.
 FT ACT_SITE 125 125 By similarity.
 FT ACT_SITE 126 126 By similarity.
 FT ACT_SITE 198 198 By similarity.
 FT SEQUENCE 362 AA; 38644 MW; 615898ADP981431 CRC64;
 SQ
 Query Match 85.3%; Score 1432.5; DB 1; Length 362;
 Best Local Similarity 84.3%; Pred. No. 5.9e-80;
 Matches 284; Conservative 22; Mismatches 30; Indels 1; Gaps 1;
 QY 1 KEVENQKLANVVIATGTTGAGASAAASNTYQAAKGVGDKLAGPELADLANVGE 60
 DB 26 KEVAQOKLSNVVIATGTTGAGASAAASNTYQAAKGVGDKLAGPELADLANVGE 85
 QY 61 QVMOJASISITDDDLKASSVAELADSDVDGIYITHTGDTLBEETAYFLNVEKTDPI 120
 DB 86 QVFOJASISITDDDLKASSVAELADSDVDGIYITHTGDTLBEETAYFLNVEKTDPI 145
 QY 121 VVVGSMRPGTANASADGMLNLYNAVAASNKDSRGKGVLTVMNDEIQSGRDVSKSINIKTE 180
 DB 146 VVVGSMRPGTANASADGMLNLYNAVAASNKDSRGKGVLTVMNDEIQSGRDVSKSINIKTE 205
 QY 181 AFKSAWGPFGMVVEKSGYWRFLPAKRTNNSFPDIKQISLPOVDIAYSIGNVTYAKA 240
 DB 206 AFKSPWGPFGMVVEKSGYWRFLPAKRTNNSFPDIKQISLPOVDIAYSIGNVTYAKA 265
 QY 241 LAONAKALLIHAGTNGSVSRVVPALQELRKNQVQIRSSR-QOQGFYLRNAEQPDDKN 299
 DB 266 LAONAKALLIHAGTNGSVSRVVPALQELRKNQVQIRSSR-QOQGFYLRNAEQPDDKN 325
 QY 300 DWVVAHDINPQKARILAMVAMTKTQDSKELQRIEMFY 336
 DB 326 DWVVAHDINPQKARILAMVAMTKTQDSKELQRIEMFY 362
 RESULT 5
 Q6PAL6 PRELIMINARY; PRT; 355 AA.
 AC Q6PAL6;
 DT 05-JUL-2004 (TREMREL. 27, Created)
 DT 05-JUL-2004 (TREMREL. 27, Last sequence update)
 DT 05-JUL-2004 (TREMREL. 27, Last annotation update)
 DE Glutaminase-asparaginase (EC 3.5.1.38) .

GN Name=aspg; OrderedLocuNames=ACIAD2088;
 OS Acinetobacter sp. (strain ADP1) .
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Moraxellaceae; Acinetobacter.
 OX NCBI_TaxID=62977;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Barbe V., Vallienet D., Fonknechten N., Kreimeyer A., Oztas S.,
 RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
 RA Ormonat L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
 RT "Unique features revealed by the genome sequence of Acinetobacter sp.
 ADP1, a versatile and naturally transformation competent bacterium." ;
 RL Nucleic Acids Res. 0:0-0(2004) .
 DR EMBL; CR543861; CAG68897.1; --
 DR GO; GO:0004067; P:asparaginase activity; IEA.
 DR GO; GO:0050417; P:glutamin-(asparagin-)ase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0006528; P:asparagine metabolism; IEA.
 DR InterPro: IPR004550; Asnase II.
 DR InterPro: IPR006034; Asp/Glutamase.
 DR Pfam: PF00710; Asparaginase; 1.
 DR PRINTS; PR00139; ASNGLNASE.
 DR ProDom: PD003221; Asp/Glutamase; 1.
 DR TIGRFAMs; TIGR00520; asnase II; 1.
 DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
 DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
 DR Complete proteome.
 KM Complete proteome.
 SQ SEQUENCE 355 AA; 37925 MW; C007PB51720C45 CRC64;
 Query Match 61.9%; Score 1040.5; DB 2; Length 355;
 Best Local Similarity 63.8%; Pred. No. 6.1e-56;
 Matches 210; Conservative 46; Mismatches 70; Indels 3; Gaps 2;
 QY 11 NVVILATGTTAGASAAASNTYQAAKGVGDKLAGPELADLANVGEQVMOJASISI 70
 DB 27 NVVILATGTTAGASAAASNTYQAAKGVGDKLAGPELADLANVGEQVMOJASISI 86
 QY 71 TNDLKLASSVAELADSDVDGIYITHTGDTLBEETAYFLNVEKTDPIVVGSMRPGT 130
 DB 87 TDKELLSARQVNDLVKPSVNGVYITHTGDTLBEETAYFLNVEKTDPIVVGSMRPGT 146
 QY 131 AMSADGMLNLYNAVAASNKDSRGKGVLTVMNDEIQSGRDVSKSINIKTEAPKASWGLP 190
 DB 147 ALSADGMLNLYNAVAASNKDSRGKGVLTVMNDEIQSGRDVSKSINIKTEAPKASWGLP 206
 QY 191 MVVEKSGYWRFLPAKRTNNSFPDIKQISLPOVDIAYSIGNVTYAKA 248
 DB 207 TLVBGKPYWFRSSVGRHTNNSFPDIKQISLPOVDIAYSIGNVTYAKA 266
 QY 249 LIHAGTNGSVSRVVPALQELRKNQVQIRSSRQOQGFYLRNAEQPDDKNVVAHDL 307
 DB 267 IINAGTNGSVGNIVPLTKKLDHDEQGIQIRSSRQOQGFYLRNAEQPDDKNVVAHDL 326
 QY 308 NPQKARILAMVAMTKTQDSKELQRIEMFY 336
 DB 327 NPQKARILAMVAMTKTQDSKELQRIEMFY 355
 RESULT 6
 ASPQ_ACIGL STANDARD; PRT; 331 AA.
 AC ASPQ_ACIGL;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Glutaminase-asparaginase (EC 3.5.1.38) (L-asparagine/L-glutamine
 DE amidohydrolase) (L-Asnase/L-Glnase) .
 GN Name=ansb;
 OS Acinetobacter glutaminasifigans.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Moraxellaceae; Acinetobacter.
 OX NCBI_TaxID=474;
 RN (1)

RP SEQUENCE.
RX MEDLINE=88243706; PubMed=3379033;
RA Tanaka S., Robinson B.A., Appella E., Miller M., Ammon H.L.,
RA Roberts J., Weber I.T., Wlodawer A.,
RT "Structure of amidohydrolases. Amino acid sequence of a glutaminase-
RT asparaginase from *Acinetobacter glutaminasificans* and preliminary
RT crystallographic data for an asparaginase from *Erwinia chrysanthemi*,"
RL J. Biol. Chem. 263:8583-8591 (1988).
RN [2]
RN SEQUENCE OF 1-60.
RX MEDLINE=78080774; PubMed=619999;
RA Holcerson J.S., Ericsson L., Roberts J.,
RT "Amino acid sequence of the diazoxonolucine binding site of
RT *Acinetobacter* and *Pseudomonas* 7A glutaminase-asparaginase enzymes,"
RL Biochemistry 17:411-417 (1978).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX PubMed=1529349;
RA Lubkowski J., Wlodawer A., Housset D., Weber I.T., Ammon H.L.,
RA Murphy K.C., Swain A.L.,
RT "Refined crystal structure of *Acinetobacter glutaminasificans*
RT glutaminase-asparaginase,"
RL Acta Crystallogr. D 50:826-832 (1994).
CC -1- CATALYTIC ACTIVITY: L-glutamine + H(2)O = L-glutamate + NH(3).
CC -1- CATALYTIC ACTIVITY: L-asparagine + H(2)O = L-aspartate + NH(3).
CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the asparaginase I family.
DR PIR, A28063; A28063.
DR PDB, 1AGX; X-ray; @=1-331.
DR InterPro; IPR004550; Asnase II.
DR InterPro; IPR006034; Asp/Glutamase.
DR Pfam; PF00710; Asparaginase; 1.
DR PRINTS; PR00139; ASNGLNASE.
DR PRODOM; PD003221; Asp/Glutamase; 1.
DR TIGRFAse; TIGR00520; asnase II; 1.
DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
KW 3D-structure; Direct protein sequencing; Hydrolase; Periplasmic.
FT ACT_SITE 12 12
FT ACT_SITE 92 92
FT ACT_SITE 93 93
FT ACT_SITE 165 165
FT BINDING 122 122
FT STRAND 3 8
FT STRAND 10 10
FT TURN 11 13
FT TURN 23 24
FT TURN 33 37
FT TURN 38 39
FT TURN 41 42
FT TURN 43 46
FT STRAND 48 52
FT STRAND 57 57
FT STRAND 59 61
FT HELIX 59 61
FT HELIX 64 78
FT TURN 79 79
FT TURN 81 82
FT STRAND 85 89
FT STRAND 92 94
FT HELIX 92 94
FT HELIX 95 105
FT STRAND 112 115
FT TURN 121 122
FT TURN 124 125
FT TURN 128 139
FT HELIX 139 140
FT TURN 140 140
FT TURN 142 143
FT TURN 145 146
FT TURN 149 153
FT TURN 154 155
FT STRAND 156 159
FT TURN 160 162

FT STRAND 164 165
FT TURN 171 172
FT STRAND 174 175
FT STRAND 182 185
FT TURN 186 187
FT STRAND 188 191
FT HELIX 200 202
FT TURN 207 209
FT STRAND 218 222
FT TURN 229 229
FT HELIX 230 236
FT TURN 237 239
FT STRAND 242 248
FT TURN 249 251
FT STRAND 252 252
FT TURN 255 256
FT HELIX 257 266
FT TURN 267 268
FT STRAND 271 276
FT STRAND 283 283
FT TURN 285 287
FT HELIX 291 294
FT TURN 295 295
FT STRAND 297 298
FT STRAND 300 300
FT TURN 301 301
FT HELIX 304 314
FT TURN 315 317
FT HELIX 321 328
FT TURN 329 331
SQ SEQUENCE 331 AA; 35485 MW; 70F1BF82389B0D31 CRC64;

Query Match 61.8%; Score 1037.5; DB 1; Length 331;
Best Local Similarity 62.0%; Pred. No. 8.6e-56;

Matches 204; Conservative 52; Mismatches 70; Indels 3; Gaps 2;

QY 11 NVVIATGTTGAGASANSATYQAAKGVNKLAGEVELDLANVRGEVWQIASSEI 70
DB 3 NVVIATGTTGAGASANSATYQAAKGVNKLAGEVELDLANVRGEVWQIASSEI 62
QY 71 TNDLLKLAASVVELADSNVDGVIYTHGDTLTETAYPLNVEKTRPIVVGSMRPGT 130
DB 63 TDKEILSLARQVNDLVKPKSVGVVITHTGDTMEETAFPLNIVHTDKRIVLVGSMRPGT 122
QY 131 AMSADGMLNVAVAASKDSRGQVLYTMNDEIQSGRDVSKSINIKTEAKSMAPLG 190
DB 123 ALSADGPLNLVSVALASNEAKNGVWVLNDSIFPAADVTKGINITHAFVSQMGALG 182
QY 191 NVVEGKSWFRLPAKRHTVNSRFDIKOI--SSLPOVDIAYSIGNVTDTPYKALAQNGAKA 248
DB 183 TIVBGPWFRRSSVKHTNNSFNIEKIQGDLPGVQIYGSDDNMPDAIYQAFKAGVKA 242
QY 249 LTHAGTNGSVSSRVVPALELR-KNGVQIIRSSRQGGFVLRNAEQPDNDVVAHDL 307
DB 243 ITHAGTNGSMANVYLVPEVRKLDHGGGLIVSSSHVAGFVLRNAEQPDNDVVAHDL 302
QY 308 NPQKRIILAMVAMTTOOSKEIQRIFWEX 336
DB 303 NPQKRIILAMVAMTTOOSKEIQRIFWEX 331

RESULT 7
ID 087079 PRELIMINARY; PRT; 354 AA.
AC 087079;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Periplasmic L-asparaginase II.
GN OrderedLocustNames-VP0374;
OS *Vibrio parahaemolyticus*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.

OX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIMD 2210633 / Serotype O3_K6;
 RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Najiima M., Nakota M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shingawa H., Hattori M., Iida T.;
 RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism
 RT distinct from that of *V. cholerae*.";
 RL Lancel 361.743-749 (2003).
 DR EMBL; AP005085; BAC61717.1; --
 DR HSSP; P06608; IHG1.
 DR GO; GO:0004067; F:asparaginase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0006528; P:asparagine metabolism; IEA.
 DR InterPro; IPR004550; Asnase_II.
 DR InterPro; IPR006034; Asp/Glutamase.
 DR Pfam; PF00710; Asparaginase_1.
 DR PRINTS; PR00139; ASNGLNASE.
 DR Prodom; PD003221; Asp/Glutamase; 1.
 DR TIGRFAMs; TIGR00520; asnase_II; 1.
 DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
 DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
 DR Complete proteome.
 SQ SEQUENCE 354 AA; 38004 MW; A1576A2EC3A6096 CRC64;
 Query Match 47.3%; Score 795; DB 2; Length 354;
 Best Local Similarity 48.3%; Pred. No. 6,8e-41;
 Matches 160; Conservative 67; Mismatches 103; Indels 2; Gaps 2;
 QY 6 CQKLVNVLATGTGTAGAGSANSATYQAKVGVDLKAGVPELADLANVGEQWQI 65
 DB 24 QSDLPNIKILATGGTITAGAGSATIS -NTAKGVGESISAVPSMTNADISGEVYSI 82
 QY 66 ASESTINDDLKLASSVAELADSNVDGIVITHTGDTLEETAYFLNLVEKTKPIYVVG 125
 DB 83 GQDMNDEWMLTAKRVNELAQDDVDGIVITHTGDTLEETAYFLDLTVKSDKPVYIYGA 142
 QY 126 MRPGTMSADGMLNTYNAAVAVNSKDSRGKGLVYTMNDEISGRDVSKEINIKTEAFSA 185
 DB 143 MRPSTMSADGCVNLYNAAVYTADESKRGVAVNMNDITFARDVTKNTTSVSTFQSP 202
 QY 186 -WGLPMVVEGKSYWFLPAKHVTNSSEFDIKQISLPQVDIAYSGVNTDTAYALAQN 244
 DB 203 NCGPIGVINSQAKQSPERKHTTETPFVDGKMLTLPVGVVYVYANASLPPVALYDA 262
 QY 245 GAKALIHAGTNGSVSRVVPALQELRKNGVQIIRSSROGGFVLRNAEQPDDKNDVVA 304
 DB 263 KEDGIVSAGVNGNLYHTVFDLEIKRASKGIVAVVSSRTPTGSTTLDAEIDDDKGVFAS 322
 QY 305 HDLNPQKRIILAMVAMTKTQDSKELQRIWEX 336
 DB 323 GTLNPQKRIILMLSLTQTKNYQDVQKMFQY 354
 RESULT 8
 QX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Hayek G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai J., Mackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck J.J., Davis N.W., Lim A., Dimalanta E.T., Potamasis K.,
 RA Apodaca R.A., Nantharavan T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RT Nature 409:529-533 (2001).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohtsuka M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsuka E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shingawa H.;
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22 (2001).
 DR EMBL; AB005526; AAC58088.1; --
 DR EMBL; AP002563; BAB37256.1; --
 DR PIR; A98108; A98108.
 DR PIR; D85953; D85953.
 DR HSSP; P0805; IJTA.
 DR HSSP; P10172; IAGX.
 DR GO; GO:0004067; F:asparaginase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0006528; P:asparagine metabolism; IEA.
 DR InterPro; IPR004550; Asnase_II.
 DR InterPro; IPR006034; Asp/Glutamase.
 DR InterPro; IPR00577; FGcy_kin.
 DR Pfam; PF00710; Asparaginase_1.
 DR PRINTS; PR00139; ASNGLNASE.
 DR Prodom; PD003221; Asp/Glutamase; 1.
 DR TIGRFAMs; TIGR00520; asnase_II; 1.
 DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
 DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
 DR PROSITE; PS00445; FGcy_KINASES_2; UNKNOWN_1.
 DR Complete proteome.
 SQ SEQUENCE 348 AA; 36850 MW; AS204EBD00910013 CRC64;
 Query Match 46.8%; Score 786; DB 2; Length 348;
 Best Local Similarity 48.3%; Pred. No. 2.4e-40;
 Matches 159; Conservative 59; Mismatches 107; Indels 4; Gaps 3;
 QY 9 LANVYTLATGTGTAGAGSANSATYQAKVGVDLKAGVPELADLANVGEQWQIAS 68
 DB 23 LPNTITLITGTTAGAGSATIS -NTAKGVENVLYNAAVPLQDIANVKGQVNNISQ 81
 QY 69 SITNDLLKLASSVAELADSNVDGIVITHTGDTLEETAYFLNLVEKTKPIYVVG 128
 DB 82 DMNDVWMLTAKKIN -TDCDKTDFVITHTGDTLEETAYFLDLTVKCDKPVVWVGAMP 139
 QY 129 GTMSADGMLNLYNAAVAVNSKDSRGKGLVYTMNDEISGRDVSKEINIKTEAFSA -WG 187
 DB 140 STMSADGPFNLVYNAVTAADKASANRGVLYVMDVTLDGRVYTKNTTDTVAETFSVNG 199
 QY 188 PLGMVVEGKSYWFLPAKHVTNSSEFDIKQISLPQVDIAYSGVNTDTAYALAQNGAK 247
 DB 200 PLGVTHNGKIDYQRPARKHTSDTPEVSKNELPKVQIVYVYANASLPAKALVDAGVD 259
 QY 248 ALIHAGTNGSVSRVVPALQELRKNGVQIIRSSROGGFVLRNAEQPDDKNDVVAHDL 307
 DB 260 GIVSAGVNGNLYKSVFTLATPAKNGAVVSSRPVPGATTQDAEVDKAGFIASGTL 319
 QY 308 NPQKRIILAMVAMTKTQDSKELQRIWEX 336
 DB 320 NPQKRIILALATQTKDPOQIQIIFQY 348
 RESULT 9
 QX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;

```
AC Q83085; Q7C019;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Periplasmic L-asparaginase II.
GN Name=anab; OrderedLocustNames=53157, SP2954;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=24577;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 24577.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; A5015308; AAN44435.1; -.
DR EMBL; A5016988; AAP18259.1; -.
DR HSSP; P00805; IJJA.
DR HSSP; P10172; IAGX.
DR GO; GO:0004067; F:asparaginase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0006528; P:asparagine metabolism; IEA.
DR InterPro; IPR004550; AsnASE_I.
DR InterPro; IPR006034; Asp/Glutamase.
DR Pfam; PR00710; Asparaginase; 1.
DR PRINTS; PR00139; ASNGINASE.
DR PRODOM; PD003221; Asp/Glutamase; 1.
DR TIGRFAMs; TIGR00520; asnase_II; 1.
DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
DR PROSITE; PS00445; FGGT_KINASES_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 348 AA; 36805 MW; BE2058CF5A010397 CRC64;

Query Match 46.6%; Score 786; DB 2; Length 348;
Best Local Similarity 48.6%; Pred. No. 2.4e-40;
Matches 160; Conservative 59; Mismatches 106; Indels 4; Gaps 3;

QY 9 LANNVLIATGCTAGAGAAANSATYQAAGVGDKLIAVPELADIANVRGSEQVMQIASE 68
DB LPNITLITAGGTIAGGSDSATKS--NTYAGKVGVENLVNAPOLKDIANVVGSEVNVIGS 81
QY 69 SITNDLLKLASSVAELDSNDVDGIYTHGTDLEETAYFLNLVEKTKPIYVVGSMRP 128
DB DANDVDVMTLAKKIN--ACDCKTDGFVITHGTDMEETAYFLDLTYKCDKPPVVMGAMRP 139
QY 82 SITNDLLKLASSVAELDSNDVDGIYTHGTDLEETAYFLNLVEKTKPIYVVGSMRP 128
DB DANDVDVMTLAKKIN--ACDCKTDGFVITHGTDMEETAYFLDLTYKCDKPPVVMGAMRP 139
QY 129 GRMASADGMLNLYNAVAVANFDSRGKGLVTNNDLISQGRDYSKINIKTEAFKSA-WG 187
DB STMSMADGPNLNAVAVTAAADKASNRGVLVNMNDVLDGRDVTKTNTTVDVATFKSVNIG 199
QY 140 STMSMADGPNLNAVAVTAAADKASNRGVLVNMNDVLDGRDVTKTNTTVDVATFKSVNIG 199
DB STMSMADGPNLNAVAVTAAADKASNRGVLVNMNDVLDGRDVTKTNTTVDVATFKSVNIG 199
QY 188 PLGMVVEGKSYMRPLPAKHRTVNSRFDIKQISLPOVDIAYSGVNTDTRAYKALAGNKA 247
DB PLGYIHNGKIDYQRTPARKRTSDTTPDVSKLNELPKVGIYVYNANMSDLPAKALVDAGYD 259
QY 200 PLGYIHNGKIDYQRTPARKRTSDTTPDVSKLNELPKVGIYVYNANMSDLPAKALVDAGYD 259
DB PLGYIHNGKIDYQRTPARKRTSDTTPDVSKLNELPKVGIYVYNANMSDLPAKALVDAGYD 259
```

```
QY 248 ALIHAGTNGSVSRVVPALQELRKNGVOIIRSSROCGGFVLNNAEOPDDKMDVVAHDL 307
DB 260 GIVSAGVGNGLYKSPFDLITRAKNGTAVASSRVPATATYODADVDKAGTFAVAGTL 319
QY 308 NPQKARILLAMVAMTKTQDSKEIQRIFMEY 336
DB 320 NPQKARVLLQLALTYTKDPQQLQIFNOY 348

RESULT 10
ID Q8CVR4 PRELIMINARY; PRT; 348 AA.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06.HI / CPT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; A5016766; AAN81991.1; -.
DR HSSP; P00805; IJJA.
DR GO; GO:0004067; F:asparaginase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0006528; P:asparagine metabolism; IEA.
DR InterPro; IPR004550; AsnASE_I.
DR InterPro; IPR006034; Asp/Glutamase.
DR Pfam; PR00710; Asparaginase; 1.
DR PRINTS; PR00139; ASNGINASE.
DR PRODOM; PD003221; Asp/Glutamase; 1.
DR TIGRFAMs; TIGR00520; asnase_II; 1.
DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 348 AA; 36881 MW; 71105BD8E007C559 CRC64;

Query Match 46.7%; Score 785; DB 2; Length 348;
Best Local Similarity 48.6%; Pred. No. 2.7e-40;
Matches 160; Conservative 58; Mismatches 107; Indels 4; Gaps 3;

QY 9 LANNVLIATGCTAGAGAAANSATYQAAGVGDKLIAVPELADIANVRGSEQVMQIASE 68
DB LPNITLITAGGTIAGGSDSATKS--NTYAGKVGVENLVNAPOLKDIANVVGSEVNVIGS 81
QY 69 SITNDLLKLASSVAELDSNDVDGIYTHGTDLEETAYFLNLVEKTKPIYVVGSMRP 128
DB DANDVDVMTLAKKIN--TDCDKTDGFVITHGTDMEETAYFLDLTYKCDKPPVVMGAMRP 139
QY 82 SITNDLLKLASSVAELDSNDVDGIYTHGTDLEETAYFLNLVEKTKPIYVVGSMRP 128
DB DANDVDVMTLAKKIN--TDCDKTDGFVITHGTDMEETAYFLDLTYKCDKPPVVMGAMRP 139
QY 129 GRMASADGMLNLYNAVAVANFDSRGKGLVTNNDLISQGRDYSKINIKTEAFKSA-WG 187
DB STMSMADGPNLNAVAVTAAADKASNRGVLVNMNDVLDGRDVTKTNTTVDVATFKSVNIG 199
QY 140 STMSMADGPNLNAVAVTAAADKASNRGVLVNMNDVLDGRDVTKTNTTVDVATFKSVNIG 199
DB STMSMADGPNLNAVAVTAAADKASNRGVLVNMNDVLDGRDVTKTNTTVDVATFKSVNIG 199
QY 188 PLGMVVEGKSYMRPLPAKHRTVNSRFDIKQISLPOVDIAYSGVNTDTRAYKALAGNKA 247
DB PLGYIHNGKIDYQRTPARKRTSDTTPDVSKLNELPKVGIYVYNANMSDLPAKALVDAGYD 259
QY 200 PLGYIHNGKIDYQRTPARKRTSDTTPDVSKLNELPKVGIYVYNANMSDLPAKALVDAGYD 259
DB PLGYIHNGKIDYQRTPARKRTSDTTPDVSKLNELPKVGIYVYNANMSDLPAKALVDAGYD 259
QY 248 ALIHAGTNGSVSRVVPALQELRKNGVOIIRSSROCGGFVLNNAEOPDDKMDVVAHDL 307
DB 260 GIVSAGVGNGLYKSPFDLITRAKNGTAVASSRVPATATYODADVDKAGTFAVAGTL 319
QY 308 NPQKARILLAMVAMTKTQDSKEIQRIFMEY 336
DB 320 NPQKARVLLQLALTYTKDPQQLQIFNOY 348
```

Db 260 GIVSAGVNGNLKYTFDTLATPAKNGTAIVASSRVPFGATTQDAEVDADKXGFVASGML 319
 QY 308 NPQKARILAMVAMTKQDSKELOIFWEX 336
 Db 320 NPOKARVILQTLTQTKDPOQIQOIFNOY 348

RESULT 11

OBAA446 PRELIMINARY; PRT; 352 AA.

AC OBAA446; 01-JUN-2003 (TRENBLREL. 24, Created)
 DT 01-JUN-2003 (TRENBLREL. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE L-asparaginase II.
 GN OrderedLocustNames=BR7257;
 OS Bacteroides thetaiotaomicron.
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OC NCBI_TaxID=818;
 (1)

SEQUENCE FROM N.A.

RC STRAIN=VPI-5482 / ATCC 29148;
 RX MEDLINE=2250858; PubMed=1263928; DOI=10.1126/science.1080029;
 RA Xu J., Bjursell M.K., Hamrod J., Deng S., Carmichael L.K.,
 RA Chiang H.C., Hooper L.V., Gordon J.I.;
 RL "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
 Science 299:2074-2076(2003).
 DR EMBL; AE016937; AA07863.1; -.
 DR HSP; P06608; 1HG1.
 DR GO; GO:0004067; P:asparaginase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0006528; P:asparagine metabolism; IEA.
 DR InterPro; IPR004550; Asnase II.
 DR InterPro; IPR006034; Asp/Glutamase.
 DR Pfam; PF00710; Asparaginase; 1.
 DR PRINTS; PR00139; ASNGLNASE.
 DR Prodom; PD003221; Asp/Glutamase; 1.
 DR TIGRFAMs; TIGR00520; asnase II; 1.
 DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
 DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
 DR Complete Proteome.
 KW SEQUENCE 352 AA; 37752 MW; 187E34EFC82BFCDF CRC64;

Query Match

Best Local Similarity 48.0%; Score 782; DB 2; Length 352;
 Matches 157; Conservative 62; Mismatches 106; Indels 2; Gaps 2;

QY 11 NVVILATGTTIAGAGASAAASATYQAAKGVDLIAGVPELADLANVGEQVQIASESI 70
 Db 27 NIIHILATGTTIAGTSSATGT-SYTAGVAIGALDLAVEIHDIANVTGEQIVRIGSQDM 85
 QY 71 TNDLLKASSVAELADSDVDGIVITHTDLEETFAVPLNVEKTRIVVGSMPRT 130
 Db 86 NDEWLTLLAKKINELLKRPDIDGIVITHTDMEETFAVPLNVTASDKVVLVGMARBT 145
 QY 131 AMSADGMLNLVAVAVASNRGKGLVTVMDIEIQSGDVGSKINIKTEAFKSA-WGPI 189
 Db 146 ALSADGPINLVAVAVTAAKESKDGVLVAMGILIGASVVKMTVDVQTQARNSGL 205
 QY 190 GNVGEGKSYWFLPAKRHTVNSBPDIKQISSLPQVDIAYSQNVDTATKALANQAKAL 249
 Db 206 GVLNKGKVCYNGITLKKTHTTQSVFDTKLTSIPKVGIVVSYGNIEDAMWTPLNNGYKGI 265
 QY 250 IHAGTNGSVSRVVPALQELAKNGVQIIRSSRQCGFLRAAEQDDKNDVVAHDLNP 309
 Db 266 IHAGVNGNTHKNIPEPSLIDARRKGIIVVRSRVPFGPTTLAEVDADKXGFVASQELNP 325
 QY 310 OKARILAMVAMTKQDSKELOIFWEX 336
 Db 326 NPOKARVILQTLTQTKDPOQIQOIFNOY 352

RESULT 12

ASG2_ECOLI STANDARD; PRT; 348 AA.

AC P00805; 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE L-asparaginase II precursor (EC 3.5.1.1) (L-asparagine amidohydrolase II) (L-Asnase II) (Colaspase).
 GN Name=aneb; OrderedLocustNames=b2957;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxID=562;
 (1)

SEQUENCE FROM N.A.

RX MEDLINE=90170867; PubMed=2407723;
 RA Jennings M.P., Beacham I.R.;
 RT "Analysis of the Escherichia coli gene encoding L-asparaginase II, aneb, and its regulation by cyclic AMP receptor and FNR proteins.";
 RL J. Bacteriol. 172:1491-1498(1990).
 (2)

SEQUENCE FROM N.A.

RC STRAIN=K12; PubMed=2144836; DOI=10.1016/0378-1119(90)90168-Q;
 RX MEDLINE=90382683; PubMed=2144836; DOI=10.1016/0378-1119(90)90168-Q;
 RA Borchon D.T.;
 RT "L-asparaginase II of Escherichia coli K-12: cloning, mapping and sequencing of the aneb gene.";
 RL Gene 91:101-105(1990).
 (3)

SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 (4)

SEQUENCE OF 23-348.

RX MEDLINE=80135739; PubMed=6766894;
 RA Malta T., Matsuda G.;
 RT "The primary structure of L-asparaginase from Escherichia coli.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 361:105-117(1980).
 (5)

PARTIAL SEQUENCE.

RX MEDLINE=80048329; PubMed=387570;
 RA Malta T., Morokuma K., Matsuda G.;
 RT "Amino acid sequences of the tryptic peptides from carboxymethylated L-asparaginase from Escherichia coli.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:1483-1495(1979).
 (6)

ACTIVE SITE.

RX MEDLINE=77140944; PubMed=321449;
 RA Peterson R.G., Richards F.F., Handschumacher R.E.;
 RT "Structure of peptide from active site region of Escherichia coli L-asparaginase.";
 RL J. Biol. Chem. 252:2072-2076(1977).
 (7)

SUBUNITS.

RX MEDLINE=73007901; PubMed=4561256;
 RA Greenquist A.C., Wriston J.C. Jr.;
 RT "Chemical evidence for identical subunits in L-asparaginase from Escherichia coli B.";
 RL Arch. Biochem. Biophys. 152:280-286(1972).
 (8)

ACTIVE SITE THR-34.

RX MEDLINE=91293312; PubMed=1906013; DOI=10.1016/0014-5793(91)80723-G;
 RA Harms E., Wehner A., Aung H.P., Roehm K.H.;
 RT "A catalytic role for threonine-12 of E. coli asparaginase II as established by site-directed mutagenesis.";
 RL FEBS Lett. 285:55-58(1991).

[9] MUTAGENESIS OF HISTIDINE RESIDUES.
RA MEDLINE=92394146; PubMed=1521538; Becham I.R., Deret C., Baet P.,
RA Wehner A., Hame E., Jennings M.P.,
RA Roehm K.H.;
RT "Site-specific mutagenesis of *Escherichia coli* asparaginase II. None
RT of the three histidine residues is required for catalysis."
RL Eur. J. Biochem. 208:475-480 (1992).
RN
RP MUTAGENESIS OF TREONINE AND SERINE RESIDUES.
RX MEDLINE=9315634; PubMed=1287659; Deret C., Hensel J., Roehm K.H.;
RA Swain A.L., Jaskolski M., Housset D., Rao J.K.M., Wlodawer A.;
RT "Crystal structure of *Escherichia coli* L-asparaginase, an enzyme used
RT in cancer therapy."
RL Proc. Natl. Acad. Sci. U.S.A. 90:1474-1478 (1993).
RN
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=93165718; PubMed=8434007; Waldner G.J., Ludkowski J., Deret C., Schleper S., Roehm K.H.,
RA Palm G.J., Wlodawer A.;
RT "A covalently bound catalytic intermediate in *Escherichia coli*
RT asparaginase: crystal structure of a Thr-89-Val mutant."
RL FEBS Lett. 390:211-216 (1996).
CC
CC -1- CATALYTIC ACTIVITY: L-asparagine + H(2)O = L-aspartate + NH(3).
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- INDUCTION: By CAMP and anaerobiosis.
CC -1- PHARMACEUTICAL: Available under the names Crastin (Bayer),
CC Elapar (Merck), Kidrolase (Rhône-Poulenc) and Lemnase (Kyowa).
CC Also available as a PEG-conjugated form (Pegaspargase) under the
CC name Oncaspar (Inzon). Used as an antineoplastic in chemotherapy.
CC Reduces the quantity of asparagine available to cancer cells.
CC -1- MISCELLANEOUS: Km = 1.15 x 10⁻⁵ M.
CC -1- MISCELLANEOUS: *E. coli* contains two L-asparaginase isoenzymes: L-
CC asparaginase I, a low-affinity enzyme located in the cytoplasm,
CC and L-asparaginase II, a high-affinity secreted enzyme.
CC -1- SIMILARITY: Belongs to the asparaginase I family.
CC -1- DATABASE: NAME=Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/ASPR/".
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC
CC -----
DR EMBL, M34277; AAA24062.1; -
DR EMBL, M34234; AAA23445.1; -
DR EMBL, U28377; AAA69124.1; -
DR EMBL, U00096; AAC75994.1; -
DR PIR, A35132; XDRC.
DR PDB, 1HO3; X-ray; A/B=23-348.
DR PDB, 1IHD; X-ray; A/C=23-348.
DR PDB, 1JAZ; X-ray; A/B=23-348.
DR PDB, 1JJA; X-ray; A/B/C/D/E/F=23-348.
DR PDB, 1NNS; X-ray; A/B=23-348.
DR PDB, 3ECA; X-ray; A/B/C/D=23-348.
DR PDB, 4ECA; X-ray; A/B/C/D=-.
DR ECHOBASE, EB0044; -
DR Ecogene, EG10046; anab.
DR InterPro, IPR004550; Anase II.
DR InterPro, IPR006034; Asp/Glutamase.
DR Pfam, PF00710; Asparaginase I.
DR PRINTS, PR00139; ASNGLNASE.

DR	ProDom: PD003221; Asp/Glutamase 1.
DR	TIGRfam: TIGR00520; aenase II; 1.
DR	PROSITE; PS00144; ASN_GLN_ASE_1; 1.
DR	PROSITE; PS00917; ASN_GLN_ASE_2; 1.
KW	3D-structure; Complete proteome; Direct protein sequencing; Hydrolase; Peptidase; Pharmacetical; Signal.
FT	SIGNAL
FT	1 22
FT	CHAIN
FT	23 348
FT	ACT_SITE
FT	34 34
FT	ACT_SITE
FT	111 111
FT	ACT_SITE
FT	112 112
FT	ACT_SITE
FT	184 184
FT	DISULFID
FT	99 127
FT	BINDING
FT	141 141
FT	MUTAGEN
FT	34 34
FT	CONFLICT
FT	49 49
FT	CONFLICT
FT	86 86
FT	CONFLICT
FT	132 132
FT	CONFLICT
FT	156 156
FT	CONFLICT
FT	171 171
FT	CONFLICT
FT	206 206
FT	CONFLICT
FT	268 268
FT	CONFLICT
FT	274 274
FT	CONFLICT
FT	285 285
FT	CONFLICT
FT	290 290
FT	CONFLICT
FT	330 330

L-asparaginase II.
 By similarity.
 By similarity.
 By similarity.
 Involved in substrate binding.
 T->A: Almost no activity.
 N -> A {in Ref. 4}.
 N -> D {in Ref. 4}.
 Missing {in Ref. 4}.
 Missing {in Ref. 4}.
 N -> D {in Ref. 4}.
 N -> D {in Ref. 4}.
 S -> T {in Ref. 4}.
 T -> D {in Ref. 4}.
 Missing {in Ref. 4}.
 Missing {in Ref. 4}.

Query Match	46.0%;	Score 772;	DB 1;	Length 348;
Best Local Similarity	47.7%;	Pred. No. 1.7e-39;		
Matches 157;	Conservative 59;	Mismatches 109;	Indels 4;	Gaps 3;

34	HELIX	36	
38	STRAND	38	
42	TURN	43	
48	STRAND	48	
54	HELIX	54	
62	HELIX	66	
67	TURN	67	
69	STRAND	78	
82	HELIX	82	
85	HELIX	97	
98	HELIX	101	
104	STRAND	108	
112	TURN	113	
114	HELIX	124	
131	STRAND	134	
140	TURN	141	
143	TURN	144	
147	HELIX	158	
159	TURN	159	
161	HELIX	163	
164	TURN	165	
168	STRAND	172	
173	TURN	174	
175	STRAND	178	
179	TURN	181	
182	STRAND	184	
186	STRAND	186	
190	TURN	191	
193	STRAND	195	

9	L	A	N	V	I	L	A	T	G	T	G	A	G	A	S	A	N	S	A	T	Q	A	K	V	G	N	D	K	L	I	G	V	P	E	L	A	D	L	A	N	V	R	G	E	O	W	Q	I	A	S	E	68						
23	L	P	N	T	I	L	A	T	G	T	I	A	G	G	D	S	A	T	K	S	-	N	I	T	V	G	K	V	E	N	L	A	N	V	A	P	O	L	A	D	I	A	N	V	K	S	Q	V	N	I	S	Q	81					
69	S	I	T	N	D	L	L	K	L	A	S	S	V	A	E	L	A	D	S	N	D	V	D	G	I	V	I	T	H	G	I	D	L	E	E	T	A	P	E	L	N	I	V	E	K	T	D	K	P	I	V	V	G	S	N	R	P	128
82	D	M	N	D	N	V	M	L	T	L	A	K	K	I	N	-	T	D	C	K	T	D	G	F	V	I	T	H	G	I	D	T	M	B	E	R	A	V	P	L	D	L	V	K	C	D	K	P	V	M	V	G	A	N	R	P	139	
129	G	T	A	S	A	D	G	M	L	N	N	A	V	A	N	S	K	D	S	R	G	K	G	V	L	Y	T	M	N	D	E	I	O	S	G	N	D	S	K	S	I	N	I	K	T	E	A	P	S	A	-	W	G	187				
140	S	T	S	N	S	A	D	G	P	F	L	N	A	V	A	V	T	A	A	D	K	A	S	A	N	R	G	V	L	V	M	N	D	T	V	L	D	G	A	D	V	T	K	T	T	D	V	A	T	E	F	S	V	N	G	199		

QY 188 PLGAVVEGKSWFRLPAKRTVNSFPDIKQISSLPQVDIANSYGVNTPDPAALAKQNGAK 247
 DB 200 PLGYTHNKIDYQRTPARKHTSDTPEFVSKLNELEPKVGVVYVYANASDLPAALADAGAD 259
 QY 248 ALIHGCTNGSSSVSRVPAALQELRKNGVOIIRSSRQGGFVLNRAEQDDKMDVAVHDL 307
 DB 260 GIVSAGVNGNLYKSVFDTLATPAKTGNAVRSVHPGATTTQDAEVDPAKGFVAGSL 319
 QY 308 NPQKARIILAMVAMTKTQDSKELOIRIFMEY 336
 DB 320 NPQKARVILQALATQTKOPQIQOIFNXY 348

RESULT 13

Q64NH4 PRELIMINARY; PRT; 352 AA.
 ID 064NH4
 AC 064NH4
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE L-asparaginase II.
 GN ORFNames=BP4215;
 OS Bacteroides fragilis.
 OC Bacterioidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=817;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YC446;
 RA Kuehara T., Yamashita A., Hiraoka H., Nakayama H., Toh H., Okada N.,
 RA Kuehara S., Hattori M., Hayashi T., Ohnishi Y.,
 RA "Genomic analysis of Bacteroides fragilis reveals extensive DNA
 RT inversions regulating cell surface adaptation."
 RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0 (2004).
 DR EMBL: AP006841; BAD50958.1; -
 SQ SEQUENCE 352 AA; 38157 MW; 3379BD95053877D1 CRC64;

Query Match 45.7%; Score 768; DB 2; Length 352;
 Best Local Similarity 47.1%; Pred. No. 3.1e-39;
 Matches 154; Conservative 60; Mismatches 111; Indels 2; Gaps 2;

QY 11 NVVILATGTTAGAGASANSATYQAAYGVKLIAGVPELADLANVGEQMOJASBI 70
 DB 27 NIHILATGTTAGTGS-ATSTNYTAGVATSTLLDAVELDIANVGEQIVRLASQDM 85
 QY 71 TNDLKLASSVAELADSDVDGIYTHGTDTLESTAYFLNVEKTDKPIVVGSMRPGT 130
 DB 86 SDEWVLIILAKKINQILKRPDIDGIYTHGTDTMEETATYFLNLYVSKNPFVLVGMRBET 145
 QY 131 AMSADGMLNLVYAVASNKDSRGKGLVTWMDIIOGRDYSKINIKTEAFKSA-WGFL 189
 DB 146 ALSADGPINLVYAVATAGAKESIGKGLVLIANGLLIAGASAIKMTIIPVQTFQAPNSGAL 205
 QY 190 GNVVSGKSYWFLPAKRTVNSFPDIKQISSLPQVDIANSYGVNTPDPAALAKQNGAK 249
 DB 206 GYIFNGKAYVYQAPLKHHTQSVFDTLATPAKTGNAVRSVHPGATTTQDAEVDPAKGFVAGSL 265
 QY 250 IHAGTNGSSSVSRVPAALQELRKNGVOIIRSSRQGGFVLNRAEQDDKMDVAVHDLNP 309
 DB 266 IHAGTNGNPFHNIIPVLELPAKKGILVVRSSRPVPGPTTMAEVDVDTQYQFIAQELNPF 325
 QY 310 QKARIILAMVAMTKTQDSKELOIRIFMEY 336
 DB 326 QKSRVILILGLTKTNDMKQIQOIFNXY 352

RESULT 14
 Q7VIE4 PRELIMINARY; PRT; 378 AA.
 ID 07VIE4
 AC 07VIE4
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE L-asparaginase (EC 3.5.1.1).
 GN Name=ansb; OrderedLocustNames=HH0662;
 OS Helicobacter hepaticus.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=32025;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 51449 / 3B1;
 RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
 RA Suerbaum S., Josephans C., Stenzenbach T., Dreescher B., Brandt P.,
 RA Bell M., Droge M., Fattmann B., Fischer H.-P., Ge Z., Hoerster A.,
 RA Holland R., Klein K., Koenig U., Macko L., Menz G.L., Nakamura G.,
 RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.,
 RT "The complete genome sequence of the carcinogenic bacterium
 RT Helicobacter hepaticus."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906 (2003).
 DR EMBL: AE017146; AAP77259.1; -
 DR HSBP; F06608; IHG1.
 DR GO; GO:0004067; F:asparaginase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0006528; P:asparagine metabolism; IEA.
 DR InterPro; IPR004550; Asnase II.
 DR InterPro; IPR006034; Asp/Glutamine.
 DR Pfam; PF00710; Asparaginase; 1.
 DR ProDom; PD003221; Asp/Glutamine; 1.
 DR TIGRFAMS; TIGR00520; asnase II; 1.
 DR PROSITE; PS00144; ASN_GLN_ASE_1; 1.
 DR PROSITE; PS00917; ASN_GLN_ASE_2; 1.
 KW Complete proteome; Hydrolase.
 SQ SEQUENCE 378 AA; 40858 MW; 45FDD24F9F9E3FAA CRC64;

Query Match 45.7%; Score 767; DB 2; Length 378;
 Best Local Similarity 47.4%; Pred. No. 3.8e-39;
 Matches 157; Conservative 71; Mismatches 99; Indels 4; Gaps 3;

QY 8 KLVNVILATGTTAGAGASANSATYQAAYGVKLIAGVPELADLANVGEQMOJASBI 67
 DB 50 KKPNIIVILATGTTAGAVDSQIKTGVNAGVISTDTLEAVPOLQEIHIQEOJIANIDS 109
 QY 68 ESTNDDLLKASSVAELADSDVDGIYTHGTDTLESTAYFLNVEKTDKPIVVGSMR 127
 DB 110 ADMNDLWTLTAQRNKLLENPKIDGIYTHGTDTMESAPFLHTTSDEKPVLTGAR 169
 QY 128 PGTASADGMLNLVYAVASNKDSRGKGLVTWMDIIOGRDYSKINIKTEAFKSA-W 186
 DB 170 PSTAISADGPINLVYAVATAGAKESIGKGLVLIANGLLIAGASAIKMTIIPVQTFQAPNSGAL 229
 QY 187 GPLGVNVEGKSYWFLPAKRTVNSFPDIKQISSLPQVDIANSYGVNTPDPAALAKQNGAK 245
 DB 230 GDMGYILDGKVFYFPKPHPTTHSEFVRSIRSLPKVDIYYSYANDGLAIAMAQALANOG 289
 QY 246 AKALIHAGTNGSSSVSRVPAALQELRKNGVOIIRSSRQGGFVLNRAEQDDKMDVAVH 305
 DB 290 TKGVLVLAGSGSIHKHKNMKLMQGLIVVQSSRLNGLIVL-ASBADKLQFISG 347
 QY 306 DLNPQKARIILAMVAMTKTQDSKELOIRIFMEY 336
 DB 348 DLNPQKARVILQALATQTKOPQIQOIFNXY 378

RESULT 15
 Q6LI34 PRELIMINARY; PRT; 354 AA.
 ID 06LI34
 AC 06LI34
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Putative periplasmic L-asparaginase II.
 GN Name=Y2787; OrderedLocustNames=PBPRB1174;
 OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Photobacterium.
 OX NCBI_TaxId=74109;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
 RA Cestaro A., Malacrida G., Simonati B., Cannata N., Bartlett D.,
 RA Valle G.;
 RT "Genome analysis of Photobacterium profundum reveals the complexity of
 RT high pressure adaptations";
 RL Submitted (Mar-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR378678; CNG3046.1; -.
 DR HSSP; P10172; IAGX.
 DR GO; GO:0004067; P:asparaginase activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0006528; P:asparagine metabolism; IEA.
 DR InterPro; IPR004550; Asnase_11.
 DR InterPro; IPR006034; Asp/Glutamase.
 DR Pfam; PF00710; Asparaginase; 1.
 DR PRINTS; PR00139; ASNGINASE.
 DR ProDom; PD003221; Asp/Glutamase; 1.
 DR TIGRPFAM; TIGR00520; asnase_11; 1.
 DR PROSITE; PS00144; ASN_GLM_ASE_1; 1.
 DR PROSITE; PS00917; ASN_GLM_ASE_2; 1.
 KM Complete proteome.
 SQ SEQUENCE 354 AA; 37904 MW; 92C77408FDC1CE32 CRC64;

Query Match 45.6%; Score 766; DB 2; Length 354;
 Best Local Simlarity 47.1%; Pred. No. 4.1e-39;
 Matches 155; Conservative 68; Mismatches 104; Indels 2; Gaps 2;

9 LANVVLATGTGTAGAGASASATYQAAKVGVDKLIAGVPELADLANVRGEQVMQIASR 68
 27 LFNKILATGTGTAGAGASATES-NYTSKVGIDELIANVPMTKADIAGEQVVKISQ 85
 69 SITNDLLKLAASSVAELADSNVDGIYITGDTTLEETAYFLMLVKTDPYIVVGSMP 128
 86 DMSDEVWLTIAVRVNALAQDDVDGVITHTGDTMGETAFLDLTVKSDKPVVLVGAMP 145
 129 GTAMGADGMLNLYNAVAVASNTDSRGKGLVTMTNDEIQSGRPVSKSINIKTEAFKA-WG 187
 146 STAMGADGPMNLVNAVVTADKDSRGVGLVAMSNVFPDARDVTKNTTSVDFQSPNFG 205
 188 PLGMVVEGKSYWFLPAKRTVNSBPDIKQISLPQVDIAVSXGNTDTAYKALQNGAK 247
 206 TIGYIHNGDVKTQKRSERKHTSDTPVDVSKLKLPRKGLVINYANASDLPRKALIDAKYD 265
 248 ALIHAGTNGSVSSRVVPAQLQELKNGVOIIRSSRQGGFVLNNAEQPDDKNDWVVAHDL 307
 266 GIVSAGVGNMNYHTIFDQLVVKASKDGTWVVRSSRTPFTSTLLDAEVDDAKGFGVASGTL 325
 308 NPQKARILAMVAMTKTQDSKELORIFWEX 336
 326 NPQKARILMLSLTQTQDYKRVQKMFQFY 354

Search completed: March 18, 2005, 21:29:40
 Job time : 183 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2005, 10:20:21 ; Search time 702 Seconds
(without alignments)
8599.100 Million cell updates/sec

Title: US-09-842-628-1

Perfect score: 1014
Sequence: 1 aaggaagtcggaacagca.....ttctctgggaactactgataa 1014

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 297611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1014	100.0	9	US-09-842-628-1
2	219	1133	9	US-09-773-260-3
3	189.2	18.7	16	US-10-406-025-1
4	145.6	11.26	10	US-09-882-227-83
5	138	13.6	17	US-10-339-670-1
6	138	13.6	18	US-10-158-865-1
7	89.4	8.8	18	US-10-425-115-4132
8	82.6	8.1	18	US-10-029-386-22613
9	78.4	7.7	18	US-10-411-910A-221
10	75.8	7.5	16	US-10-029-386-20582
11	74.6	7.4	17	US-10-283-122A-14804

Result No.	Score	Query Match Length	ID	Description
12	72.2	7.1	960	US-10-472-928-4159
13	72.2	7.1	2162598	US-10-472-928-4979
14	70	6.9	2364	US-10-473-687-3
15	70	6.9	2718	US-10-738-986-15
16	69.4	6.8	1749	US-10-411-910A-218
17	69.2	6.8	1605	US-10-259-165-297
18	69	6.8	1314	US-10-156-761-1793
19	69	6.8	1758	US-10-411-910A-209
20	69	6.8	9025608	US-10-156-761-1
21	68.6	6.8	779	US-10-029-386-20935
22	68.6	6.8	985	US-10-363-345A-31407
23	68.6	6.8	985	US-10-363-345A-31408
24	67.2	6.6	791	US-10-029-386-20494
25	67.2	6.6	828	US-10-029-386-25443
26	67	6.6	1512	US-10-425-115-150486
27	66.8	6.6	1859	US-10-425-115-150483
28	66.4	6.5	1725	US-10-411-910A-196
29	66.4	6.5	9025608	US-10-156-761-1
30	65.4	6.4	1444	US-09-070-927A-56
31	65.4	6.4	2475	US-10-767-701-12575
32	65.4	6.4	2475	US-10-260-238-5598
33	65.4	6.4	2475	US-10-425-114-16619
34	65.4	6.4	64492	US-10-378-083-1
35	65.2	6.4	871	US-10-029-386-20422
36	65.2	6.4	1089	US-10-156-761-6834
37	65.2	6.4	1761	US-10-156-761-3138
38	65	6.4	1398	US-10-369-493-31934
39	65	6.4	1485	US-10-411-910A-267
40	64.6	6.4	786	US-10-187-267A-32
41	64.6	6.4	36321	US-10-187-267A-1
42	64.4	6.4	1268	US-10-716-803-2
43	64.4	6.4	5392	US-10-716-803-1
44	64	6.3	1107	US-10-156-761-2234
45	64	6.3	1312	US-10-437-963-29396

ALIGNMENTS

RESULT 1

US-09-842-628-1

Sequence 1, Application US/09842628

Patent No. US2002064862A1

GENERAL INFORMATION:

APPLICANT: ROBERTS, JOSEPH

APPLICANT: MACALLISTER, THOMAS W.

APPLICANT: SETHURAMAN, NATARAJAN

APPLICANT: FREEMAN, ABBIE G.

TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND ITS USE IN

TITLE OF INVENTION: ANTIVIRAL AND ANTICANCER THERAPY

FILE REFERENCE: 023032/0108

CURRENT APPLICATION NUMBER: US/09/842,628

CURRENT FILING DATE: 2001-04-27

PRIOR APPLICATION NUMBER: 08/050,482

PRIOR FILING DATE: 1995-04-25

PRIOR APPLICATION NUMBER: PCT/US92/10421

PRIOR FILING DATE: 1992-12-04

PRIOR APPLICATION NUMBER: DR P 4140003.8

PRIOR FILING DATE: 1991-12-04

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 1014

TYPE: DNA

ORGANISM: Pseudomonas sp.

FEATURES:

NAME/KEY: CDS

LOCATION: (1)..(1008)

US-09-842-628-1

Query Match 100.0%; Score 1014; DB 9; Length 1014;

Best Local Similarity 100.0%; Pred. No. 4,16-259;

Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	AGGAACTGAGAAACCAAGCAAGAACTGGCCAAAGTGTGATATCTGGCCACACGGACGGACAC	6
Db	1	AAGGAATGGAAGAACCAAGCAAGAAAGCTGGCCAAAGTGTGATATCTGGCCACACGGACGGACAC	60
OY	61	ATGCGCCGACCTGAGCGCCAGACGGCGCCAAACAGCGCCACCTACCAAGCTGCAAGGTTGGC	12
Db	61	ATGCGCCGACCTGAGCGCCAGACGGCGCCAAACAGCGCCACCTACCAAGCTGCAAGGTTGGC	12
OY	121	GTGCAACAAGTGAATTGGCCGCGTGGCCGAGCTGGCCAGACTGGCCAAATGTGGCCGCGAG	18
Db	121	GTGCAACAAGTGAATTGGCCGCGTGGCCGAGCTGGCCAGACTGGCCAAATGTGGCCGCGAG	18
OY	181	CAGGTGATGCAAGATCGCCTCCGAAAGCATCAACAAGACGACTGTCAAGCTGGCAAGC	24
Db	181	CAGGTGATGCAAGATCGCCTCCGAAAGCATCAACAAGACGACTGTCAAGCTGGCAAGC	24
OY	241	AGCGTGGCCGAGCTGGCCGACAGCAATGACGTGATGAGCATGTCTACACCATGAC	30
Db	241	AGCGTGGCCGAGCTGGCCGACAGCAATGACGTGATGAGCATGTCTACACCATGAC	30
OY	301	GACACCCCTGGAAAGAAACCGGCTACTTTTGAACCTGTGAAAGAACGCAAGCGCATC	36
Db	301	GACACCCCTGGAAAGAAACCGGCTACTTTTGAACCTGTGAAAGAACGCAAGCGCATC	36
OY	361	GTGCTGTGCTGGTTTCATACGCGCCCGGACACGCGCATGTCCGCGACGCAATGCTCAACCTG	42
Db	361	GTGCTGTGCTGGTTTCATACGCGCCCGGACACGCGCATGTCCGCGACGCAATGCTCAACCTG	42
OY	421	TACAAAGCCCTGGCCGTGGCCGACAAACAGAACTCGCGCGGACAGAGGCGTGTGTGACC	48
Db	421	TACAAAGCCCTGGCCGTGGCCGACAAACAGAACTCGCGCGGACAGAGGCGTGTGTGACC	48
OY	481	ATGAAGACACAGATCCACGTCGCGGCGTGAAGTGAAGAGTGCATCAATCAAGACCGAA	54
Db	481	ATGAAGACACAGATCCACGTCGCGGCGTGAAGTGAAGAGTGCATCAATCAAGACCGAA	54
OY	541	GCCTTCAAGACCGCTTGGGCGCCGCTGGGCGATGTTGTGAAAGGCAAGTGTGTGTTCC	60
Db	541	GCCTTCAAGACCGCTTGGGCGCCGCTGGGCGATGTTGTGAAAGGCAAGTGTGTGTTCC	60
OY	601	CGCTGCGCGGCGCAAGCGGCCACAGGCTCAATCCGAGTTTGGACATCAAGCAATCAGCAGC	66
Db	601	CGCTGCGCGGCGCAAGCGGCCACAGGCTCAATCCGAGTTTGGACATCAAGCAATCAGCAGC	66
OY	661	CTGCCCCAGGTGACATATGCTACAGCTATGCGCAAGTACCGACACGCGCTTACAAAGGCC	72
Db	661	CTGCCCCAGGTGACATATGCTACAGCTATGCGCAAGTACCGACACGCGCTTACAAAGGCC	72
OY	721	CTGGCACAAGAACGGCGCCAAAGCGCTGATCATGCGCGCACCGGCAATGGCTCGGTTGCG	78
Db	721	CTGGCACAAGAACGGCGCCAAAGCGCTGATCATGCGCGCACCGGCAATGGCTCGGTTGCG	78
OY	781	TGCGGGGTGTGGTCCAGCGCTGCAAGAGAGTGGCGGAAGAAAGCGGCTGGAGATCAATGGTTCC	84
Db	781	TGCGGGGTGTGGTCCAGCGCTGCAAGAGAGTGGCGGAAGAAAGCGGCTGGAGATCAATGGTTCC	84
OY	841	TCAAGTCAACAGGGCGGTTTCTGTGCTGTGATGCGGACAGCGCCGACACAAGAACACAC	90
Db	841	TCAAGTCAACAGGGCGGTTTCTGTGCTGTGATGCGGACAGCGCCGACACAAGAACACAC	90
OY	901	TGGGTGTGGTCCCAAGCTTGAACCGCAGAGAGGCCCGCATCTTGGCGATGTTGGCAATG	96
Db	901	TGGGTGTGGTCCCAAGCTTGAACCGCAGAGAGGCCCGCATCTTGGCGATGTTGGCAATG	96
OY	961	ACCAAGACCAAGACAGCAAGAGAGAGCTGGACAGGAGATTTCTGGGAATACAGATTA	1014
Db	961	ACCAAGACCAAGACAGCAAGAGAGAGCTGGACAGGAGATTTCTGGGAATACAGATTA	1014

RESULT 2
US-09-773-260-3
; Sequence 3, Application US/09773260

```

: Patent No. US20020102251A1
: GENERAL INFORMATION:
: APPLICANT: Donald L Darden
: TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES
: ASPARAGINASE IN THE TREATMENT OF HUMAN
: HEMATOLOGIC AND AUTOIMMUNE DISEASE
:
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
:
: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: PASCSEQ for Windows 2.0
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/773,260
: FILING DATE: 31-Jan-2001
: CLASSIFICATION: <Unknown>
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/094,435
: FILING DATE: 1998-06-09
: APPLICATION NUMBER: <Unknown>
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 234/274
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
:
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1133 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: SEQUENCE DESCRIPTION: SEQ ID NO: 3:
:
: US-09-773-260-3
:
: Query Match 21.6%; Score 219; DB 9; Length 1133;
: Best Local Similarity 52.8%; Pred. No. 3.7e-48;
: Matches 521; Conservative 0; Mismatches 460; Indels 6; Gaps 2.
:
: QY CCAAGCTGTGATCTGTGGCCACCGGCGGCACCATCGCCGGCGCTGGCGCCAGCGCGGCCA 88
: DB 128 CCGAAGTACATCTTATGCGCACGAGAGGCACATCGCTTGGGGGAATCTAGCGTCA 187
:
: QY ACAGCGCCACCTTACCAAGCTGCGCAAGTTGGCGCTGCACAGCTGATTTGCCGCGTGGCG 148
: DB 188 AGAG---TAGTACTTGTCTGGAGACAGTCAACGTTGATTAAGCTTCTTGCAACCGTCCCTG 244
:
: QY AGCTGCCGCACTTGGCCCAATGTGGCGCGCGGCGAGAGTGCAATGCGCTCCGGAAGCA 208
: DB 245 CCAATCAAGCATGACCAACCATCAAGGAGTGAAAGATCTCAAGCATTTGGCTCCCAAGAGA 304
:
: QY TCACCAACGACGACCTGCTCAAGCTGGCGAAGACAGCTGGCCGACGACGAATG 268
: DB 305 TGAAGGGGTAAAGTGTGGCTTAAACTTAAGCCAAAGCTGTCAATGAGTCTCTCCGCCAAAAG 364
:
: QY ACGTCGATGGCATGCTATCAACCATGAGCGACCGACACCCCTGAAAGAAACCGCTACTTTT 328
: DB 365 AGACCGAAGCGGTATCATCAACCATGGAATCTGACACCATGGAAGAGACCGGTTTCTTCC 424
:
: QY TGAACCTGTGTGAAAAGACGACCAAGCGATGCTGTGTGTGGTTTCATATGGGCCCCGGGCA 388

```

Db 425 TCNACCTCACGGTGAAAGCCAAAACCTGCTCTCTTAGGCGCCATGCGTCCAGGCT 484
 Qy 389 CGGCCAATGTCGGCGCGGCGATGCTCACTGTAACAACCGCGCGCGCGCGCGCAACA 448
 Db 485 CTTCAATGAGTGTGATGCGCCCATGAACTCTATTAACCGCGTAAATGATGATCAACA 544
 Qy 449 AGGACTCGCGCGCAAGGCGTGTGTCATGAAACGACGATCGATCGCGCGCTG 508
 Db 545 AAGCCTCTAATAAAGAGGTGTGATGTGATGAAACGATGATTAACGCGCGCAAG 604
 Qy 509 AGTGAACGATGTCATCAACATCAAGACCGAACCTTTAA--GAGCGCTGCGCGCG 565
 Db 605 AAGCGACCAAGCTCAACACCGCATCATGTCATTTGCTTGGCCCAACAGGTAA 664
 Qy 566 TGGGCAATGATGGAAGGCAAGTGTGATGTCGTTCCGCTGCGCGCAACCGCAACCG 625
 Db 665 TCGGCAATGATGTAATGGAAGTGTGATGTCATGTCATGTCATGTCATGTCATGTCAT 724
 Qy 626 TCNACTCGGATTCGATCAACAGCAGATGACAGCTGCGCGCGCGCGCGCGCGCGCG 685
 Db 725 TTGCAAGTGAATGATTAATGCAAAATGGAAGACTCCCGAGATGATTAATTTTAA 784
 Qy 686 GCTATGCGCAAGTCAACGCGCGCTACAGGCGCTGCGCAAGAACGCGCGCAAGCGCG 745
 Db 785 CTCAACCCGATGATCTGATGTTTATGTCATGACCGCTTACGCGCGCAAGAA 844
 Qy 746 TGAATCATGCG 805
 Db 845 TCATCATGCG 904
 Qy 806 ACCTGCGCAAGAACG 865
 Db 905 AAGCAGCGCAATCG 964
 Qy 866 TCGTGAACG 925
 Db 965 CCGAAGAGGCTGAAGGTGATGAAGAACTTGTTTGTGCTACAGAGATCTCAAC 1024
 Qy 926 CGCAGAGGCG 985
 Db 1025 CTCGAAAGCGCGAGGCTCTTATGTTAGCCCTCAACAACTAGTGAAGAGCGCA 1084
 Qy 986 TCGAGCGCATTTTCTGGAATACTGAT 1012
 Db 1085 TCCAAAAGATCTTCTCCACTATTAAT 1111

RESULT 3
 US-10-406-025-1
 / Sequence 1, Application US/10406025
 / Publication No. US20030186380A1
 / GENERAL INFORMATION:
 / APPLICANT: No. US20030186380A1ozymes Biotech, Inc.
 / APPLICANT: Thomas, Michael D.
 / APPLICANT: Sloma, Alan
 / TITLE OF INVENTION: Methode for producing secreted polypeptides having L-asparaginase
 / TITLE OF INVENTION: activity
 / FILE REFERENCE: 10289.200-US
 / CURRENT APPLICATION NUMBER: US/10/406, 025
 / PRIOR FILING DATE: 2003-04-01
 / PRIOR APPLICATION NUMBER: US 60/369,192
 / NUMBER OF SEQ ID NOS: 8
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 1
 / LENGTH: 1128
 / TYPE: DNA
 / ORGANISM: Bacillus subtilis
 US-10-406-025-1

Query Match 18.7%; Score 189.2; DB 16; Length 1128;
 Best Local Similarity 50.2%; Pred. No. 2.9e-40;
 Matches 498; Conservative 0; Mismatches 488; Indels 6; Gaps 1;

Qy 20 AAGAGCTGGCGCAACGATGATCTGCGCACCGCGCGCGCAATCGCGCGCGCTGCGCGCA 79
 Db 143 AGGATCTGCGCAACATTAATTTTATGCGACAGAGGCGCGATGCTGCTGCGCATCAAT 202
 Qy 80 GCGCGCGCAACAGCGCGCATTAACAGCTGCGCAAGGTTGCGCTGCAACAGCTGATTCGG 139
 Db 203 CGAAATCTCAACATGATTAATTAAGCAGGTGTTGCGCGCTGGAATCACTGATGAG 262
 Qy 140 GCGTCCGAGCTGCGCGCATGCGCGCATGTCGCGCGCGCGCGCGCGCGCGCGCGCG 199
 Db 263 CAGTTCCAGAAATGAGCAATTCGCAAGCTGACGCGCGAGAGATGTTAAGTGTGCGCA 322
 Qy 200 CGAAAGCATCAACAGCGCATGCTGCAAGCTGCGCAAGCGCGCGCGCGCGCGCGCGCG 259
 Db 323 GCAAAATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 382
 Qy 260 ACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 319
 Db 383 CTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 442
 Qy 320 CTTATTTTGAATCTGCG 379
 Db 443 CTTATTTTGAATCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 502
 Qy 380 GCG 439
 Db 503 GACCTTCAACCG 562
 Qy 440 CCAGCAACAGGATCTGCG 499
 Db 563 CAGGTGCG 622
 Qy 500 CCGGCG 559
 Db 623 CAGCG 682
 Qy 560 GCG 619
 Db 683 TGGGCTTCTGCG 742
 Qy 620 ACAGCG 679
 Db 743 ATACGAGGACAGGATTTCTGCGTTTCTAATCTTGAATGATGATGATGATGATGATGAT 802
 Qy 680 CTTACGCTATGCG 739
 Db 803 TCTATGATACCAAAATGACGAGAGCTACCTGTTGACGCTGCTGTAAGCGCGAGCAA 862
 Qy 740 AGCG 799
 Db 863 AGGGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 922
 Qy 800 TCGAGAGCTGCG 859
 Db 923 CGGACG 982
 Qy 860 TCGTCTGCG 919
 Db 983 TCGTCAACCAAAAC-----CAAGACTATGCGGAAAGGACTTGTGCGCATGAACTCTT 1036
 Qy 920 TGAACCG 979
 Db 1037 TAAACCG 1096
 Qy 980 AGAGCTGCG 1011
 Db 1097 AAAAAATCAAGCTTATTTCAATGATTAATGA 1128

RESULT 4
 US-09-882-227-83
 / Sequence 83, Application US/09882227

Publication No. US20030158396A1
 GENERAL INFORMATION:
 APPLICANT: Kleantous, Harold
 APPLICANT: Al-Garawi, Amal
 APPLICANT: Miller, Charles
 APPLICANT: Tomb, Jean-Francois
 APPLICANT: Oomen, Raymond P.
 TITLE OF INVENTION: Identification of Polynucleotides
 TITLE OF INVENTION: Encoding No. US20030158396A1el Helicobacter Polypeptides in the
 FILE REFERENCE: 06132/047002
 CURRENT APPLICATION NUMBER: US/09/882,227
 PRIOR FILING DATE: 2001-06-15
 PRIOR APPLICATION NUMBER: US 08/902,615
 PRIOR FILING DATE: 1997-07-29
 NUMBER OF SEQ ID NOS: 638
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 83
 LENGTH: 1196
 TYPE: DNA
 ORGANISM: Helicobacter pylori
 FEATURES:
 NAME/KEY: CDS
 LOCATION: (41)...(1132)
 US-09-882-227-83

Query Match 14.4% Score 145.6; DB 10; Length 1196;
 Best Local Similarity 50.5%; Pred. No. 1,le-28;
 Matches 436; Conservative 0; Mismatches 419; Indels 9; Gaps 3;

1 AAGGAAGTGGAGAACGACGAGAAAGTGGCCAACTGCTGATCTTGGCCACCGCGGCAAC 60
 131 AAGGGGAGGTTATGCTCAAAATTTACCAATGCTTTACTGGCGACAGGGGAGGACG 190
 61 ATCGCGCGGCTGGCG 120
 191 ATTGCAAGGAGTGGCG 247
 121 GTGCAAGAGCTGATGCG 180
 248 ATCAAGAGCTTTGAGAGCTATCTTATGCTTTAAGAGCTGCTGCGATTTAAGGGAG 307
 181 CAGGTATGAGATGCTGCTCCGAAAGCATCACCAAGCAAGCAAGCTGCTCAAGCTGGCAAGC 240
 308 CAGATTTCTAATGATGCGCTCACAGACATGAATGAAGAGTATGTTCAAGCTCCGCAAA 367
 241 AGCGTGGCGGACGCTGGCG 300
 368 CGTGCCCAAGATTGCTATGATATGCGCTATTTCAAGGCGTGTATCAAGCTGCGACG 427
 301 GACACCTGGAAGAAACGCTACTTTTGAACCTGCTGGAAGAAACGCGACAGCGGATC 360
 428 GACATTTTGAAGAGAGCGCGCTATTTTAACTTAACTTAACTTAACTTAACTTAACTTAA 487
 361 GTGTGTGCTGTTCCATGCG 420
 488 GTGTGTGCTGCGATGCGTATGCTGCTTTTGAAGCGCGGATGCGGCTTTGAATTA 547
 421 TACAAAGCG 480
 548 TATAATGCTGTAGGCTAGCGCTCATGTAAAGAGTGAATTAAGGCGCTGTATGTGTG 607
 481 ATGAAGACGAGATCGATCG 540
 608 ATGAGAGATTAATTTTATGCGCTAGAGAGATTAATTAATTAATTAATTAATTAATTAAT 667
 541 GCGTTCAAGAGC---GCTGGGCGCGCGCTGGGCGATGCTGTGGAAGCAAGTGTGAG 597
 668 ACCTTTAAAGCCTTAATATAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 727
 598 TTCCGCTGCG 654
 728 TACATGACGCGCTTTGAGAAACACACACAGAGAGGAAATTTTCCCTTCAAGACTCAAA 787

655 AGACGCTGCCCGAGTGGACATGCTTACAGTATGCAACGTCACGACAGCGGCTTAC 714
 788 ACCCGCTGCCCGTAAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 847
 715 AAGCGCTGGACAGACAGCG 774
 848 CAAGGAGCTTAACTCGCATGCAAAAGCGTGTGATAGCGCGGCGGTAAATGGAAT 907
 775 GTGTGTGCGCGGCGGTGTCGACGCGCTGAGAGAGCTGCGCAAGACGCGTGCATCAT 834
 908 GTAGCGCTGGGCTTTTAAAGCGATGCAAAAGAGAGCGCAAAATGCGGCTTATGT 967
 835 GCTTGTGACGTCACAGCGCGGT 858
 968 CTTCTAGCAGGCTAATATACCGGT 991

RESULT 5 US-10-329-670-1

Sequence 1, Application US/10329670
 Publication No. US20040018503A1

GENERAL INFORMATION:
 APPLICANT: Fleischmann et al.
 TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fragm
 FILE REFERENCE: PB186P1
 CURRENT APPLICATION NUMBER: US/10/329,670
 PRIOR FILING DATE: 2002-12-24
 PRIOR APPLICATION NUMBER: US 09/643,990
 PRIOR FILING DATE: 2000-08-23
 PRIOR APPLICATION NUMBER: US 08/487,429
 PRIOR FILING DATE: 1995-06-07
 PRIOR APPLICATION NUMBER: US 08/426,787
 PRIOR FILING DATE: 1995-04-21
 NUMBER OF SEQ ID NOS: 1
 SOFTWARE: Patent version 3.1
 SEQ ID NO 1

LENGTH: 1830121
 TYPE: DNA
 ORGANISM: Haemophilus influenzae

FEATURE:
 NAME/KEY: misc feature
 LOCATION: (4747)..(4747)
 OTHER INFORMATION: n equals a, t, g or c

FEATURE:
 NAME/KEY: misc feature
 LOCATION: (9921)..(9921)
 OTHER INFORMATION: n equals a, t, g or c

FEATURE:
 NAME/KEY: misc feature
 LOCATION: (10150)..(10150)
 OTHER INFORMATION: n equals a, t, g or c

FEATURE:
 NAME/KEY: misc feature
 LOCATION: (29298)..(29298)
 OTHER INFORMATION: n equals a, t, g or c

FEATURE:
 NAME/KEY: misc feature
 LOCATION: (36543)..(36543)
 OTHER INFORMATION: n equals a, t, g or c

FEATURE:
 NAME/KEY: misc feature
 LOCATION: (36551)..(36551)
 OTHER INFORMATION: n equals a, t, g or c

FEATURE:
 NAME/KEY: misc feature
 LOCATION: (36636)..(36636)
 OTHER INFORMATION: n equals a, t, g or c
 NAME/KEY: misc feature
 LOCATION: (40808)..(40810)
 OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

LOCATION: (152530)..(152530)

Query Match 13.6%; Score 138; DB 17; Length 1830121;
Best Local Similarity 48.5%; Pred. No. 2,7e-26;
Matches 474; Conservative 0; Mismatches 491; Indels 12; Gaps 3;

QY 40 ATCTGGCCACCGGCGGACCAATCGCCGCTGGCGCCGCGGCGCAACAGCCGAC 99
DB 802735 ATCTGGCAACGGGGGTACCACTTCAGAGAGCGGCAAAAGTTCCGTAAA---TTCTCG 802791
QY 100 TACCAAGCTGCCAAGTTGGCGTGCACAAAGCTATTTGGCGGCGGCGGCGGCGGCGG 159
DB 802792 TATTAAGCTGACCAATTAAGTATTAATCTTTAATTAAGCTGACCAAGAAATGAAAT 802851
QY 160 CTGGCCCAATGTGCGCGGCGGAGAGGTGATGACATGCTGCTCCGAAGATCAACCAAGAC 219
DB 802852 ATTGCAACATTTAAAGGTGAGAAATTTGTAATAATAGTTTCAAGACATGAAATACGA 802911
QY 220 GAACCTGCTCAAGCTGGCAAGACGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 279
DB 802912 GTCTGGCTAAACCTGGCAAAAGCCATCATGCTCAATGTAAGTAAAGTAACT-----GATGGA 802965
QY 280 ATCTGATCAACCCATGACGACGACCCCTGGAAGAAACCGCTACTTTTGAACCTCGTG 339
DB 802966 TTGTGATTAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803025
QY 340 GAAAAGACCAACCAAGCCGATCGTGGTGGTTCATGCGCCCGGCGGCGGCGGCGGCGGCGG 399
DB 803026 GTAAATATGTAAAAACCGGTGTTCTGTTGGGGGCAATCGCTTCCGCAACGAAAAAGT 803085
QY 400 GCCGACGCGATGCTCAACCTGTACACGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGG 459
DB 803086 GCTGTGCGCCCATTAATCTTTAATCATGCTGCTGTGCGGCGGCGGCGGCGGCGGCGG 803145
QY 460 GGCAAGGCGCTGCTGTGATCAATGACGAGATTCAGATCCGCGGCGGCGGCGGCGGCGG 519
DB 803146 GGTCTGTGTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 803205
QY 520 TCGATCAATCAATCAACCAAGCCCTTCAAGAGCGC---CTGGGGCGCGCTGGCGATGCTG 576
DB 803206 ACGATGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 803265
QY 577 GTGAGGCGAATGCTGATGCTGCTGCGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGG 636
DB 803266 CATACACCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803325
QY 637 TTTCGATCAATCAATCAACCAAGCCCTTCAAGAGCGC---CTGGGGCGCGCTGGCGATGCTG 696
DB 803326 TTTAACGTAGAAAAATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 803385
QY 697 GTCAACGACCAAGCGGCTTACAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 756
DB 803386 GCACTGTGCAACCAATTAAGCAATTAATGCTGATGATGATGATGATGATGATGATGAT 803445
QY 757 GGCAACGCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 816
DB 803446 GGAAGTGGCAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803505
QY 817 AACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 876
DB 803506 GATAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 803565
QY 877 GAGCAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 936
DB 803566 GAACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803625
QY 937 CGCATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 996
DB 803626 CGCGGCTCTTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803685
QY 997 TTCTGGGAATGATGAT 1013
DB 803686 TTGGAAGACTTCTAAGA 803702

RESULT 6
US-10-158-865-1
Sequence 1, Application US/10158865
Publication No. US20040203093A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fragm
FILE REFERENCE: P8186P2C1D1
CURRENT APPLICATION NUMBER: US/10/158,865
PRIOR FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 09/557,884
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 08/476,102
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (40810)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g

FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:

NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match 13.6%; Score 138; DB 18; Length 1830121;
Best Local Similarity 48.5%; Pred. No. 2,7e-26;
Matches 474; Conservative 0; Mismatches 491; Indels 12; Gaps 3;

QY 40 ATCTGGCCACCGCGGCAATCGCCGCGCTGGCGCGCGGCGCAACGCCGACC 99
DB 802735 ATCTGGCCACCGCGGCAATCGCCGCGCTGGCGCGCGGCGCAACGCCGACC 99
QY 100 TACGAGGCTGCCAAGGTGGCGGTGCAAGCTGATGGCGGCGGAGCTGGCGGAC 159
DB 802792 TATTAAGCTGCACATTAAGTATTAATCTTAATTAAGCTGTACGAAATGAATAAT 802851
QY 160 CTGGCCAAATGTGGCGGCGGAGGAGGTGATGCAGATCGCTCCGAAGCATCCACGAC 219

```

Db 802852 ATTGCAACATTAAAGTAGAGCAAAATTGTAATAAGTTCAACAAGCATGATGACGA 802911
Oy 220 GACCTGCTCAAGCTGGCAAGCAGCTGGCCGAGTGGCCGACAGCAATGACCTGCATGAC 279
Db 802912 GTCTGGCTTAAACTGGCAAAAGCCATCAATGCTCAATGTAATAAGTACT-----GATGA 802965
Oy 280 ATCTGATCAACCATGCGACCGACACCTGGAGAAACCCGCTACTTTTGAACCTCGTG 339
Db 802966 TTTGTCATTACCCATGATGATACATGAGAAAGACGGCTTATTTCTAGATTAAAC 803025
Oy 340 GAAAAGACCGCAACACCGATGCTGCTGCTTCAATGCGCCCGGACCGCCGATGCTCC 399
Db 803026 GTAAATATGTAATAAACCGGTTGTTCTGTTGGGCAATGCGCTGCAACAGAAAAGT 803085
Oy 400 GCCGACGCGATGCTCAACCTGTACAAACGCGCTGGCCGCTGGCCAGCAAGGATCGCGC 459
Db 803086 GCTGATGGCCCATTAATCTTTACAAATGTCGTGTCGACGACAGCAAAAATCAAGT 803145
Oy 460 GGCAAGGCGCTGCTGCTGACCATGAACGACGATCCAGTCGCGGCGTGAACGTGACGAG 519
Db 803146 GGTCTGCTGTTTATGTCGCAATGATATGATGATGATGATGATGATGATGATGATGAT 803205
Oy 520 TCGATCAATCAACACCGAAGCTTCAAGAGCGC---CTGGGGCCCGCTGGCGATGGTG 576
Db 803206 ACCAGTACGACCGCAGTGAACGTTCCATTCAACAAATTAATGTTCTCTAGGCTATAT 803265
Oy 577 GTGAAGGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 636
Db 803266 CATACAGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803325
Oy 637 TTGCAATCAACAGATCAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Db 803326 TTTAATGTAATAAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 803385
Oy 697 GTCAACGCAACGCGCTTACAGGCGCTGCGACAGAAAGCGGCAAGCGCTGATTCATGCG 756
Db 803386 GACCTGTCGAACCAATTAAGCGATTAATCAATGCTGCTGCTGCTGCTGCTGCTGCTG 803445
Oy 757 GGCACCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 816
Db 803446 GGAATGTCGAATGAAATGTAATGCTGCACTTAATGCTTGAATAAAGCCGCAAAA 803505
Oy 817 AACGCGTCGACAGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
Db 803506 GATAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 803565
Oy 877 GAGCAGCCCGCAACGAAGACGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 936
Db 803566 GAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 803625
Oy 937 CGCATCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996
Db 803626 CGCGTCTCTTGCATTAAGCTTAACTAAAGTCCAAAGTAAATTCACAAATAT 803685
Oy 997 TTCTGGGAATGCTGATA 1013
Db 803686 TTGGAAGACTTCTAAGA 803702

```

```

RESULT 7
US-10-425-115-43132
; Sequence 43132, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28

```

```

; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 43132
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MHT4577_139344C.1
US-10-425-115-43132

```

```

Query Match      8.8%; Score 89.4; DB 18; Length 437;
Best Local Similarity 53.9%; Pred. No. 7.9e-14;
Matches 228; Conservative 0; Mismatches 191; Indels 4; Gaps 2;

```

```

Oy 317 CCAGCTTCTTTTGAACCTGCTGGAAGAACCGACAGCCGATGCTGCTGCTGCTGCTGCTGCT 376
Db 3 CGGCTTCTTCTGATGATGACGAGGAGTGCAGAACCTCTGCGCTGGGGGGGCA 62
Oy 377 TGCCCGCCCGGACCGCCATGTCGCCCGGACCGCATGCTCAACCTGTACAAACCGCTGGGCG 436
Db 63 TGCTACCGGCACTGCGATGAGCGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 122
Oy 437 TGCCAGCAACAGACTCGCGCGCAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 496
Db 123 CAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 182
Oy 497 AGTCGGGCTGACGCTGACGTAAGTGAATCAATCAACAGCCGATTC---AAGAGCG 553
Db 183 TGATGAGGCTGACGACCAAGCAACATTAATGCGCGCAGACTTTCATACAGCA 242
Oy 554 CCGGGGCGCGCTGGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
Db 243 AGTTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
Oy 613 AAGCGCCACAGCTCACTCCGAGTTCGATCAATCAAGCATGACAGCTGCGCCAGGTG 672
Db 303 GTAAGCATACCGTGTATCTACTCACTGACGCTGACGAGCTGGAAGTCCGAAAGTTC 362
Oy 673 GACATGCTTACAGCTATGGAACGTCACGACACGCGCTCAAGAGCCCTGGCACAGAAC 732
Db 363 GGCATGCTTACACTATGACACGACGCTGATGCGCCGACCAAGAGTCTGATGACGAA 422
Oy 733 GGC 735
Db 423 GGC 425

```

```

RESULT 8
US-10-029-386-22613/C
; Sequence 22613, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: ABOICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22613
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008103.27
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.87
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: P28481, EVALU8 7.00e-12

```

OTHER INFORMATION: EST HUMAN HIT: AL110383.1, EVALUE 0.006+00
 OTHER INFORMATION: NT HIT: g14756466, EVALUE 0.006+00
 US-10-029-386-22613

Query Match 8.1%; Score 82.6; DB 16; Length 990;
 Best Local Similarity 47.2%; Pred. No. 5,6e-12;
 Matches 364; Conservative 0; Mismatches 389; Indels 18; Gaps 3;

```

QY 5 AAGTGGAGACACGAGAGAGAGCTGGCCAACTGGTATCTCTGGCCACCGGCGGACCATCG 64
DB ACGAGAGATCTGTACACGATCTGCTAACAGAGCGCGGACAGAGGACATCGCCAAAGAG 871
QY 65 CCGGCGCTGGCGCCAGCGCGGACCAAGCGCCACTACAGGCTTCCAAAGTTGGCTCG 124
DB ACGCGCGCCAGGCAATCGCCAGAGAGAGCGCGCCCAAGGAATCGCCAGAGAGAGCGCG 811
QY 125 ACAAAGTATTTGCGCGCTGCGGAGCTGGCGGACCTTGGCCAAATGTGCGCGGAGAG 184
DB CCCACGCGATCGCCAGAGAGAGCGCGCCAGAGGATCGCCAGAGAGAGCGCGCGCCAG 751
QY 185 TGAT-----GCAATGCGCTCTCCGAAAGATCAACAGAGAGAGCTGCTCAAGCTGGCA 238
DB GCATTCGACAGAGAGAGAGCGCGCCAGAGGATCGCCAAAGAGAGAGCGCGCCAG----- 697
QY 239 GCAAGCTGGCGAGCTGGCGCCAGCAAGCAATGAGCTGATGGCATCTGCTCAATCCATGGCA 298
DB GATGCGCCAAAGAGAGAGAGCGCGCGCCAGAGGATCGCCAAAGAGAGAGCGCGCCAGCGCA 640
QY 299 CCGACACCTTGGAGAGAAACCGCTTACTTTTGAACCTGTGAAAGAGAGAGAGAGAG 358
DB TCGCCAAAC---GAGGATCTCGCGCCAGAGATCGCAGAGAGAGAGAGAGAGAGAGAGAG 583
QY 359 TCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 418
DB CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 523
QY 419 TGTACAAAGCGGTGGCGGTGGCGCAAGCAATGCGCGCGAGAGAGAGAGAGAGAGAGAG 478
DB AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 463
QY 479 CCAATGAACGAGAGAGATCAAGTCCGCGCGGTGACGTGAGCAAGTCAATCAAGAGAGCG 538
DB CCGCCCAAGGAGATCGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 403
QY 539 AAGCTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 598
DB AGGAGATGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 343
QY 599 TCGGCTGCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658
DB TCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 283
QY 659 GCGTGGCCAGAGTGAATCGCTTCAAGTATGAGCAAGTCAAGAGAGAGAGAGAGAGAG 718
DB AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 223
QY 719 CCTTGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 765
DB ACGCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 172

```

RESULT 9
 US-10-411-910A-221

Sequence 221, Application US/10411910A
 Publication No. US20040209256A1
 GENERAL INFORMATION:
 APPLICANT: Dillon, Harrison F.
 TITLE OF INVENTION: Methods and Compositions for Evolving Hydrogenase Genes
 FILE REFERENCE: H2041203-P
 CURRENT APPLICATION NUMBER: US/10/411,910A
 NUMBER OF SEQ ID NOS: 343
 SOFTWARE: PatentIn version 3.2

SEQ ID NO 221
 LENGTH: 1518
 TYPE: DNA
 ORGANISM: Chlamydomonas reinhardtii
 US-10-411-910A-221

Query Match 7.7%; Score 78.4; DB 18; Length 1518;
 Best Local Similarity 46.3%; Pred. No. 7,6e-11;
 Matches 339; Conservative 0; Mismatches 381; Indels 12; Gaps 2;

```

QY 46 GCCACCGCGGACCATTCGCGCGCTGGCGCGCAAGCGGCGCAAGCGCCACTTACAG 105
DB GCCCGCGGACAGACCATTCGCGCGCGCGCGCAAGCGGCGCGCAAGCGCCACTTACAG 219
QY 106 GCTGCAAGTTGGGTGAGCAAGCTGATTTGCGCGCGCGCGCGAGAGCTG-----GCCAG 159
DB CTGGCGCTGAGAGAGAGCTGAGCAAGCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAG 279
QY 160 CTGGCAATGTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 219
DB GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 339
QY 220 GACTGTCTCAAGCTGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 279
DB AGCCCGCGCAAGCTGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
QY 280 ATCTGATCAATCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 339
DB CTGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 459
QY 340 GAAAAGACCGGCAAGCGCGAGCTGCGGTGCGGTTCATGCGCGCGCGCGCGCGCGCGCGCG 399
DB GAGAGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 519
QY 400 GCGGAGCGGATGCTCAACCTGTGACCAAGCGCGGTGCGCGCGCGCGCGCGCGCGCGCG 459
DB CCGGCTGAGTGGCGCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 579
QY 460 GCGAAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 519
DB TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 639
QY 520 TCGATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
DB GGCATTCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 699
QY 574 GTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 633
DB GAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 759
QY 634 GAGTTGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 693
DB ATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819
QY 694 AAGCTTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 753
DB CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879
QY 754 GCCGCGACCGCG 765
DB ACCACCGCGCGCG 891

```

RESULT 10

US-10-029-386-20582/c
 Sequence 20582, Application US/10029386
 Publication No. US20030194704A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanel, David K.
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
 EXPRESSION ANALYSIS TWO

SEQ ID NO 3
LENGTH: 2364
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1) .. (2364)
OTHER INFORMATION: Maize optimized v1p3B
US-10-473-687-3

Query Match 6.9%; Score 70; DB 18; Length 2364;
Best Local Similarity 45.5%; Pred. No. 1.4e-08;
Matches 327; Conservative 0; Mismatches 385; Indels 6; Gaps 2;

13 AACGACGAAAGCTGCGCAACGCTGATCTGGCCACCGCGGACCATCGCCGCGCT 72
Db ATCATGAAAGGACCTCTGACAAAGAGAGAGAGTTCCGGTGAACATCTCCGAC 975
73 GCGCCGAGCGCGCCAAAGCGCCACTTACAGGCTGCCAGGTTGGCGTGCACAAAGCTG 132
Db CTCTCCAAACCTTCTCCAAACCGAACTACGCAAGGCGCAAGGCTCCAAACGAGACGCG 1035
133 ATTGCGGCGTCCGAGCTGGCCCACTGGCCCAATGTGCGGCGGAGAGAGTGAATGACG 192
Db AAGATCATGTGAGAGGCGCAAGCGCGCTACGCTCTGTGGGCTTCAAGATGTCACAGAC 1095
133 ATCGCTCCGAAAGATGACCAACGACGACTGTCTCAAGCTGGCAAGCGCTGGCCGAG 252
Db TCCATCACCGTGTCTAAGGCTTACGAGCGCAAGCTCTACAGAGATCAACAGGTGACAA 1155
253 CTGGCGGACAGCATGACGTGATG--CATGTGATCACTCATGGACCGACACCTG 309
Db GACTCCCTCTCGAGATGCTGTACGCGGACATGACCAAGCTCTGCGCCGAGACAGTCC 1215
310 GAAAGAAACCGCTACTTTTGAACCTCTGTGAAAAACCGCAACCGCATCTCGTGTGTC 369
Db GAGCAATCTTACTTACCAACCAACATCGCTTCCGAAACGATGATGATCAACAGATC 1275
370 GATTCCATGCGCGCGCGGACCGCCATGTCCGCGGACGCGATCTCAACTGTACACGCC 429
Db ACCTTCAACCAAGAAATGAATCTCTCCGCTACGAGGCCACCGCACTTTCACATCTC 1335
430 GTGGCGGTGGCCAGCAACAGGACTCGCGCGGCAAGGCGGTGCTGTGACATGAACGAC 489
Db TCCACCGGCGGACATGACCTTCAACAAACCAAGGTGCTCTCGAGGCGCAGATATCTC 1395
430 GAGATC---CACTCCGCGGCGTACGTGAGCAAGTGAATCAATCAAGCCGAAACCTTC 546
Db ACCCTCTCGGCTTCCACCGACGCGCGGTATCATGCGCTGCGATCATCTCCAGACCTTC 1455
547 AAGACGCGCTGGGCGCGCGTGGCATGTGTGGAAGCAAGTCTGTACTGTTCGCGCTG 606
Db CTCACCCCGCATTAAGGCTTGGCATGTGTGGAAGCAATCCAGTCTGTGAACCTTC 1515
607 CCGGCGAAGCGCCACACGCTCACTCGAGTTGACATCAAGCAATCAAGCAAGCTTGCCC 666
Db ACCTCAAGTCTTACCTCGCGAGGCTCTCGCCACGACCTCTCCAAACAGAGAAC 1575
667 CAGGTGACATGCTCAAGCTATGGAACGTCACCGCAACGCGCTTCAAGGCGCTTGG 724
Db AAGTCATGTGCGCGCGCATGCGCTTCAATCTCCAAATGTGGAAGACGCAACCTGG 1633

RESULT 15

US-10-738-986-15
Sequence 15, Application US/10738986
Publication No. US20040209241A1
GENERAL INFORMATION:
APPLICANT: Hermanson, Gary
APPLICANT: Geall, Andrew
APPLICANT: Mloeh, Mary Kopke
TITLE OF INVENTION: Codon-Optimized Polynucleotide-Based Vaccines Against Human
CYTOMEGALOVIRUS INFECTION

FILE REFERENCE: 1530.0580001
CURRENT APPLICATION NUMBER: US/10/738.986
CURRENT FILING DATE: 2003-12-19
PRIOR APPLICATION NUMBER: US 60/435,549
PRIOR FILING DATE: 2002-12-23
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn version 3.2
SEQ ID NO 15
LENGTH: 2718
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: codon-optimized sequence of hCMV gB
US-10-738-986-15

Query Match 6.9%; Score 70; DB 18; Length 2718;
Best Local Similarity 44.8%; Pred. No. 1.4e-08;
Matches 357; Conservative 0; Mismatches 430; Indels 9; Gaps 2;

15 CCAGCAAGCTGGCCAAAGTGTGATCTGGCCACCGCGGACCATCGCCGCGCTGG 74
Db CAGCTGGGCGAGGACAAAGATCTGTGGGCAACCAAGAGCCAGAGTGCACACT 1835
75 GCGCAGCGCGCCAAAGCGCCACTTACAGAGCTGCCAGGTTGGCTTCACACTGAT 134
Db GCCCAGCTTGAAGATCTTATCGCGGCAACAGCGCTTACAGATAGTGAATCTGTT 1895
135 TGCCGCGGTGCGGAGCTGGCGGACCTGGCCAAATGTGCGCGGAGAGATGACAGT 194
Db CAAGAGATGATCACTGAGAGCATGACACCGTGAACAGATATCGCCCTGACAT 1955
195 CGCTCCGAAAGATGACCAACGACGACTGTCAAGCTGCGCAAGCGTGGCGGACT 254
Db GACCCCTTGAGAAACACCGACTTACAGGCTGTGAAGCTGTA---CAGCCAGAAAGACT 2012
255 GCGCAGCAGATGACGTGATGAGCATGTATCAACCATGAGCAACCTTGAAGA 314
Db GAGGAGCAGCAACGTGTGACCTGAGAGAGATCATGAGGAGTTCAACAGTACAGCA 2072
315 AACCGCTACTTTTGAACCTGTGAAAAACCGCAACGCGATGTGTGCTGTTTC 374
Db GAGGTGAAGTACTGTGAGAGCAAGGTGTGACCCCTGCCCCCTTACATGAAGGCTT 2132
375 CATGCGCCCGGACCGCATGTCTCGCGGACGCACTGTCAACCTGTACACGCGTGGC 434
Db GAGCAGCTGATGAGCGCGCTGGCGCGCGCGCAAGGCGGTGGCGCATCGGCGC 2192
435 CGTGCGCAGCAACAGATCTGCGCGGCAAGGCGTGTGTGACCAATGAAGAGAT 494
Db CGTGGCGCGCGCGGTGACCGGTGTGAGAGGCGGTGCGCACTTCTGAAGAACCTCTT 2252
495 CCAATCGCGGCGGTGACGTGACCAAGTCAATCAAGACCGAAGCGCTTCAAGAGCGC 554
Db CGGCGCTTCAACATCTCTGTGTGACATGCGCGTGTGATCATCACTACTGATCTA 2312
555 CTGGGCGCGGTGGCATGTGTGGAAGCAAGTGTGATCTGCTTCCGCTTCCGCGCA 614
Db CACCAAGGAGAGAGGTGTGACCAAGCCCTGGAAGACCTGTTCCCTACTGTGAG 2372
615 GCGCAGCAGGTCACTCGAGTTGACATCAAGAGATCAAGAGCTGCCCGCAGTGA 674
Db CGCCAGCGGACCAACCGTGAACAGCGGAGACCAAGGACCAAGCGCTGCG-----AGGC 2426
2373 CATGCGCTAAGCTTATGCAAGTCAACGACGCGCTTACAAAGGCGCTTGCACAAAGCG 734
Db CCCCCAGCTACAGAGAGAGCTGTACAAAGCGGAGAGAGGCGCCCGGCCCCCAG 2486
735 GCGCAGCGCTGTATCATGCGCGGACCGGCAATGCTCGTGTGTCGCGGCGGTGCG 794
Db CAGGAGCGCAGCAACCGCGCGCGCGCTTACCAACCAAGAGCGGCTTACAGATCTGCT 2546
795 AGCTTGCAGAGCTG 810

Db 2547 GGCCCTGGCCAGGCTG 2562

Search completed: March 22, 2005, 13:09:36
Job time : 713 secs

This Page Blank (uspto)

FEATURES	source
Db	635 CATGGCGGGGGCGGGGGCGGCAAGCGCCGACGCGCCACTCAACCGGCGCCGAAGTGGC 576
Qy	120 CGTGGACAACTGATTGCGCGGCGTGGCCGAGCTGGCCGACCTGGCCAAATGTGGCGGCGGA 179
Db	575 GGTGGACCAAGTTGGTGGCGGACGTAACCCAGCTCAAGGACATTCGCAACGTCGCGGGGGA 516
Qy	180 GCAGGTGATGAGATTCGGCTCCGAAAGCATGCACCAACGACGCTGCTCAAGCTGGCAAG 239
Db	515 ACAAGTGTTCAGATCGCTTCGGAAAGCTTCAACCAACGAAAGCTGCTGGAACTGGGGA 456
Qy	240 CAGGTGGCCGAGCTGGCCGACAGCAATGACGTGATGGCATTCGATCAACCCATGGCAC 299
Db	455 GACCGTGGCCAAAGCTGGCCGATAGGACGAGCGTCAACGCGCATTCGATCAACCCACGGTAC 396
Qy	300 CGACACCCCTGGAAAGAAACCGGCTACTTTTGAAACCTCGTGGAAAAAGACCGAACGGCAT 359
Db	395 CGAACCCCTGGAAAGAACCGGCTACTTCTTGACCTGTGTGAGCACACCGAAGAACCTTAT 336
Qy	360 CGTGTGTGCGGTTTCATGCGCCCGCGGACCGGCAATGTCGCGCGGACGCGATGCTCAACT 419
Db	335 CGTGTGTGCGGCTCGATGCGCCCGGGGACCGCGATGTCGCGCGGACGCGCATGCTCAACT 276
Qy	420 GTACAAAGCCCGTGGCCCGTGGCGGACGAACAAGACTCTGGCGCGGCAAGGCGCTGTGTAC 479
Db	275 GTACAAAGCCCGTGGCGGCGTGGCGGACGAAGTCCGACGCGGCAAGGGCGTGTGATAC 216
Qy	480 CATGAACGACGAGATCCAGTCCGAGCCGTGAAGTGAAGCAAGTGCATCAACATCAAGACCGA 539
Db	215 CATGAACGACGAGATTCCTCTCGCGCGGGAAGCGGACGAAGATGCTCAACATCAAGACCGA 156
Qy	540 AGCCTTCAAGAGCGGCTGGGCGCGGCTGGGCAATGCTGTGAAGGCAAGTGTACTGTT 599
Db	155 AGCCTTCAAGAGCCCGTGGGCGCGGCTGGGCAATGCTGTGAAGGCAAGTGTACTGTT 96
Qy	600 CCGGCTGGCGG 610
Db	95 CCGGCGACCGG 85
RESULT 2	B2571362
LOCUS	B2571362
DEFINITION	meb2_1851.x1 meb Pseudomonas aeruginosa genomic clone meb2_1851,
ACCESSION	B2571362
VERSION	B2571362.1
KEYWORDS	GI:27206423
SOURCE	GSS.
ORGANISM	Pseudomonas aeruginosa Pseudomonas aeruginosa Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
REFERENCE	1 (bases 1 to 871)
AUTHORS	Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE	Whole-Genome-Sequence Variation among multiple isolates of Pseudomonas aeruginosa library
JOURNAL	J. Bacteriol. (2002) In press
COMMENT	Contact: Chris K. Raymond Genome Center University of Washington Box 352145, Seattle, WA 98105-2145, USA Tel: 2062216954 Fax: 2066857244 Email: craymond@u.washington.edu Classes: shotgun. Location/Qualifiers 1..871

```

/clone_lib="mah"
/note="Environmental isolate. Whole genomic shotgun
library."

```

Query Match	37.9%	Score 384.2;	DB 8;	Length 871;
Best Local Similarity	77.7%	Pred. No. 2e-71;		
Matches 491; Conservative	0;	Mismatches 134;	Indels 7;	Gaps 2;

OY	377	TGGGCCCCGGCAGCGGCATGCGCGCGACGGCACTGCTCAACCTGTACAACGCCGGGCGG	436
Db	34	TGGGTACCGAGCCCCCCTCGAGGTCCAGGTAATGCAATAGCTGATATTCGAATTT---CC	90
OY	437	TGGCCAGCAACAAGAACTCGCGCGGCAAGGCGTGCTGTGTGACCAATGAACGACGATCC	496
Db	91	TGCAGCCCGATTAAGTCGGCAACGCGGCAAGGCGTGTGATCATCATGAACGACGATCC	150
OY	497	AGTCCGGGCGGTGANTGTAGCAATGCATTCACATCAAGATCCGAAGCCTTCAAGAGCGCT	556
Db	151	TCTCCGCGCCGAGCGCAGCAAGATGTGTCAACATCAAGACCGAAGCCGTGCAAGAGCCCGT	210
OY	557	GGGGCCGCGTGGGCATGTGGTGGAAAGCAAGTGTGTCTGGTTCGCGCTGCAGCAAGC	616
Db	211	GGGGCCGCGTGGGGATGTGTGTTCAGAGGCAAGAGCTACTGTGTTCCGCGCACCGGTGAAGC	270
OY	617	GCCACACGCTCACTCCGAGTTGCACTCAAGCAGATCAGACGCTTGCCTGCCAGGTGACA	676
Db	271	GGCACACGTTGAATCCGAGTTTCAACATCAAGCAGATCTTTCGCCCTGCTCCGTTGGAAA	330
OY	677	TGGCTTACAGTTATGGCAAGTCAACGACACGCGCTTCAAGGACCCTTGACACAGAACGGCG	736
Db	331	TGGCTTACAGTTACGTTAAGTCAAGCCACACCGCTTCAAGGACCCTTGACACAGAACGGCG	390
OY	737	CCAAAGCGCTGATCCATGCGCGCACCGGCAATGGCTCGTGTCCGTCCGCGGTGTGCGAG	796
Db	391	CCAAAGGCGATCATCATGCGCGGACCGGCAATCGGTTCCGTTGCCCCGCGCGTGGGCGCGC	450
OY	797	CCCTGCAGAGCTGCGCAAGAACGGCGTGCAGATCATTC---GTGTGTCACGTCAACAG	852
Db	451	CTTGTGTGGAATCTGCGCAACAGGCGGTGACGATTAATTCGGGTCTTGCAAGTCAATGCC	510
OY	853	GGCGGTTTCGTGCTGCGTTAAACGCCGACAGGCCGACGACAAAGACGACTGGGTGATGCC	912
Db	511	GGCGGTTTCGTCTGCGCAACGCCGACAGCGCGACGACAAAGACGACTGGATGTGTGCC	570
OY	913	CACGACCTGAACCCGCAAGAGGCCCGATCCTGTGCATGTGTGGCAATGACCAAGACCCAG	972
Db	571	CACGACCTGAACCCGCAAGAGGCCCGATCCTGTGCGTGTGCGGATGACCAAGACCAAT	630
OY	973	GACAGCAAGAGCTGCAAGCGCATTTTCTGGGA	1004
Db	631	GACATGATGAGCTGATCGGATTTTTTGGAA	662

RESULT 3	
BZ570599/c	
LOCUS	BZ570599 768 bp DNA linear GSS 17-DEC-2002
DEFINITION	msl2_144c.y2 msh Pseudomonas aeruginosa genomic clone msh2_144c,
	genomic survey sequence.
ACCESSION	BZ570599
VERSION	BZ570599.1 GI:27205660
KEYWORDS	GSS.
SOURCE	Pseudomonas aeruginosa
ORGANISM	Pseudomonas aeruginosa

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 766)	Spencer, D.H., Raymond, C.K., Smith, E.B., Sims, E.B., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.	Whole-Genome-sequence variation among multiple isolates of <i>Pseudomonas aeruginosa</i>	Library J. Bacteriol. (2002) In press	Contact: Chris K. Raymond

Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 20622216954
Fax: 2066857244
Email: cgraymond@u.washington.edu
Class: shotgun.

FEATURES

source
1.768
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="mh2_1446"
/clone_1lb="mh"
/note="Environmental isolate. Whole genomic shotgun library."

ORIGIN

Query Match 36.0%; Score 364.6; DB 8; Length 768;
Best Local Similarity 82.0%; Pred. No. 2.9e-67;
Matches 443; Conservative 0; Mismatches 95; Indels 2; Gaps 2;

241 AGCGTGGCCGAGCTGGCCGACAGCAATGACGTGATGCGATGCTCATCCCATGGCACC 300
604 ACCGTGCTTACCTGGCCGATAGCGAAACAGTGGCGGATGATGATCACCCCGGTACC 545
301 GACACCTCTGAGAAACCGCTTACTTTTGAACCTCTGTGAAAAGACGACAGACCCGATC 360
544 GACACCTCTGAGAAACCGCTTACTTTTGAACCTCTGTGAAAAGACGACAGACCCGATC 486
361 GTCGTGCTGCTTCAATGCGCCCGCCGACGCGCATGTCGCGCATGCTCAACCTG 420
485 GTGTGTGCTGCTTCAATGCGCCCGCCGACGCGCATGTCGCGCATGCTCAACCTG 426
421 TACAACGCGCTGCGCGCGCGCAACAAGACCTGCGCGCAAGGCGCTGTGTGAC 480
425 TACAACGCGCGCGCGCGCGCGCAACAAGACCTGCGCGCAAGGCGCTGTGTGAC 367
481 ATGAACGACGAGATTCAGTCCGCGCGCGTGAAGTGAAGACGATGATCAACATCAAGACGAA 540
366 ATGAACGACGAGATTCAGTCCGCGCGCGTGAAGTGAAGACGATGATCAACATCAAGACGAA 307
541 GCGTTCAGAAAGCGCTCTGCG 600
306 GCGTTCAGAAAGCGCTCTGCG 247
601 GCGTTCAGAAAGCGCTCTGCG 660
246 GCGTTCAGAAAGCGCTCTGCG 187
661 CTGCCCCAGGTGAGCAATGCGCTTACAGCTTATGCGCAACGCGACGCGCTTACAGAGCC 720
186 CTGCGCTCGGTGGAATCGCTTACAGCTTATGCGCAACGCGACGCGCTTACAGAGGCC 127
721 CTGCGCAAGAAAGCG 780
126 CTGCG 67

RESULT 4
BZ571432/c 840 bp DNA linear GSS 17-DEC-2002
LOCUS mh2_1883.x1 msh Pseudomonas aeruginosa genomic clone mh2_1883,
DEFINITION genomic survey sequence.
ACCESSION BZ571432
VERSION BZ571432.1 GI:27206493
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 840)

AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.B., Sims, E.B., Haastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-genome-sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 20622216954
Fax: 2066857244
Email: cgraymond@u.washington.edu
Class: shotgun.

FEATURES

source
1.840
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="mh2_1883"
/clone_1lb="mh"
/note="Environmental isolate. Whole genomic shotgun library."

ORIGIN

Query Match 18.8%; Score 191; DB 8; Length 840;
Best Local Similarity 75.9%; Pred. No. 2.7e-30;
Matches 236; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

30 CAGGTGTGATCTCTGCGCACCGCGCGCACATCGCGCGCTGCGCGCGCGCGCGCA 89
407 CAGGCGGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 348
90 CAGGCGCGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 149
347 CAGGCGCGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 288
150 GCTGCGCGCGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 209
287 GCTGCGCGCGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 228
210 CACCAACGACGACCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 269
227 CACCAACGACGACCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 168
270 CGTGATGCGATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 329
167 CGTGATGCGATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 108
330 GAACTCTGCTG 340
107 GAACTCTGCTG 97

RESULT 5
BZ571326/c 802 bp DNA linear GSS 17-DEC-2002
LOCUS mh2_1835.x1 msh Pseudomonas aeruginosa genomic clone mh2_1835,
DEFINITION genomic survey sequence.
ACCESSION BZ571326
VERSION BZ571326.1 GI:27206387
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCES 1 (bases 1 to 802)
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.B., Sims, E.B., Haastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-genome-sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond

Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: Shotgun, 4/20/2004

FEATURES

```

1. 802
  /organism="Pseudomonas aeruginosa"
  /mol_type="genomic DNA"
  /strain="MSH"
  /db_xref="taxon:287"
  /clone="mh2.1835"
  /clone_1fb="mh"
  /notes="Environmental isolate. Whole genomic shotgun
  library."

```

ORIGIN

	Query Match	Similarity	DB 8;	Score 139;	Length 802;
	Best Local	Similarity 71.0%;	Pred. No. 3, 1e-19;		
	Matches 184;	Conservative 0;	Mismatches 75;	Indels 0;	Gaps 0
QY	82	GCGCACACAGCGCCACTTACAGGCTGCGCAAGGTTGGCGTCGACACAGCTGATGGCCGC	141		
Db	320	GCGCACACACCGGCTCTCCACACCCCGCAAGGTCGCCGCTGACACAGTTGCTGCACGA	261		
QY	142	GTGCGAGAGCTGCGCCGACCTGGCCATGTCGCGCGCGACAGAGTATGACAGATTCGCTCC	201		
Db	260	GTACCGAGCTTCAGAGGGATCGTATCGTCCGCGCGACAGAGTGTCCAGATTCGCTCG	201		
QY	202	GAAGGATCAACCAACGACGACTGTCTCAAGTGGCGAAGCGGTGGCCGAGCTGGCCGAC	261		
Db	200	GATGAGCTTACCATCGAAGAACTGTGGATCTGGGCGAAGACCGTACCGACTGTGCCAT	141		
QY	262	AGCATGACGTGACATGGCATGTCATCAACCCATGGCACCGACACCCCTGGAGAAACCGCC	321		
Db	140	AGCGACGACGTGACAGCGCATGGTATGTCGTCTCGGTACGACACACGTGGAGAGACCGCC	81		
QY	322	TACTTTTGAACCTCGTGG	340		
Db	80	TCTTCTCTACCCCTGGGGG	62		

RESULT 6

LOCUS	644 bp	DNA	linear	GSS 12-SEP-1998
nbhd0008p12r	CUGI	Rice BAC library	<i>Oryza sativa</i> [japonica	
cultivar-group)	genomic clone nbhd0008p12r,	genomic survey		
	sequence.			

ACCESSION	AO157105
VERSION	AO157105.1
KEYWORDS	GI:3554130
SOURCE	GS8.
ORGANISM	<i>Oryza sativa</i> (japonica cultivar-group)
	<i>Oryza sativa</i> (japonica cultivar-group)
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Eriocaulaceae; Oryzaceae; Oryza.
REFERENCE	1 (bases 1 to 644)
AUTHORS	Wing, R. A. and Dean, R. A.
TITLE	A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL	Unpublished (1998)
COMMENT	Contact: Wing RA

FEATURES

Source

1. .644
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="mbx0008p12r"
/tissue_type="leaf"
/lab_host="E. coli DH10B"
/clone_lib="CGI Rice BAC Library"
/note="Vector: pBeloBAC1; Site: 1: HindIII, Site: 2:
HindIII; Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Armuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9 %. Two high density filters,
each containing 16,432 clones (doubly spotted), represent
the whole library for colony screening."

ORIGIN

Query Match	Best Local Similarity	5.0%; Score 91.4; DB 8; Length 644;
Matches 224; Conservative	0; Mismatches 221; Indels	0; Gaps 0;
Qy	175	GGCGAGCAGGTGATGTCAGATCGCCTCCGAAAGCATCACCGACGACCTGCTCAAGCTG 234
Db	57	GTCATGTAATTCATGAGCGGTGGGAGTTTGCACTTCGCTCGAAGACGCTTTGGCTTC 116
Qy	235	GCAAGCAGCGGTGGCCGAGCTGGCCGACAGCAATGACGTGATGGCATCGTCATCAACCAT 294
Db	117	GCCCGACGAGCAGMAAGACCATCTGCGCGAACCCCAAGATCGACGGGTGTCGTGACCGCAC 176
Qy	295	GGCACCGACACCTGTGAAGAAACGGCTACTTTTGAACCTCGTGAAGAACCGACAG 354
Db	177	GGAACCGACACGATGAGAGAGACGGCTTAATATGGCCGATCTTTCATCGATTAGGAGAG 236
Qy	355	CCGATCGTCGTGATCGGTTCCATGCGGCCCGGACCGGCATGTCGCGCGACGGATGCTC 414
Db	237	CCGGTCGTCTTTAACGGGTGCGACGGCGGACCGGATACCGGATGCGAGTGTCCACGC 296
Qy	415	AACCTGTACACGCGCGTGGCCGTGGCCGACGAAACAGGACTCGCGCGGCAAGGGCGTCTG 474
Db	297	AACATTCGCCAGCGGGTTCGCGATGCTGACATCTGAGACCGCGCGGTGCGGGGCTCTT 356
Qy	475	GTCACCATGAAACGACGAGATCCAGTCCGGGCGTGAAGCTGAGCGAAGTGCATCAATCAAG 534
Db	357	TTGTGTTTCAGACGAGGAGTTTACGACGCCCGGACGATGACGAAGTCCACACTTCGCG 416
Qy	535	ACCGAAGCCTTTCAGAGACGCGCTGGGGGCCCTGGGCGCATGATGTGGAAGCAAGTCGTAC 594
Db	417	ACCGACACGTTCAATTTCTGCCGAACATGGGAGCTGGGTGAAGTCGATGAGTCACTGATC 476
Qy	595	TGTTTCGCGCTGCGGCGCAAGCGCC 619
Db	477	ATCGTGACCCGCGCCCGGTGGCC 501

RESULT 7

BH376509		DNA	GSS 10-DEC-2002
LOCUS	757 bp		
DEFINITION	AG-ND-137P15.TF.1 ND-TM Anopheles gambiae genomic clone		
ACCESSION	AG-ND-137P15, genomic survey sequence.		
	BH376509		

ACCESSIO


```

VERSION      BH376509.1  GI:17322651
KEYWORDS
SOURCE
ORGANISM      Anopheles gambiae (African malaria mosquito)
               Anopheles gambiae
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
               Anopheles.
REFERENCE     1 (bases 1 to 757)
               Hong, Y.S., Hogan, J.R., Wang, X., Sarkar, A., Sim, C., Loftus, B.J.,
               Ren, C., Huff, E.R., Carlile, J.L., Black, K., Zhang, H.-B.,
               Gardner, M.J. and Collins, F.H.
               Construction of a BAC library and generation of BAC end
               sequence-tagged connectors for genome sequencing of the African
               malaria mosquito Anopheles gambiae
JOURNAL      Mol. Genet. Genomics 268 (6), 720-728 (2003)
MEDLINE      22542063
PUBMED       12655398
COMMENT       Other GSSs: AG-ND-137P15.TR.1
               Contact: Brendan J Loftus
               Department of Eukaryotic Genomics
               The Institute for Genomic Research
               9712 Medical Center Dr., Rockville, MD 20850, USA
               Tel: 301 838 0208
               Fax: 301 838 3543
               Email: b.loftus@igr.org
               This clone is from an A. gambiae BAC library (ND-TAM) provided by
               P.H. Collins and sequenced by The Institute for Genomic Research
               (TIGR). The BAC library was generated from A. gambiae PEST strain
               DNA. All DNA was extracted from newly hatched first instar larvae
               to minimize the inclusion of DNA from microorganisms that inhabit
               the gut. The DNA is derived from mixed sexes of larvae. The BAC
               library was constructed at Texas A&M University BAC Center
               University, College Station, Texas 77843-2123, USA using a HindIII
               partial digest.
               Seq primer: M13 For
               Class: BAC ends.
FEATURES
source       Location/Qualifiers
               1..757
               /organism="Anopheles gambiae"
               /mol_type="genomic DNA"
               /strain="PEST"
               /db_xref="taxon:7165"
               /clone="AG-ND-137P15"
               /clone_1lb="ND-TAM"
               /note="Vector: pECBAC1; Site_1: HindIII"
ORIGIN
Query Match      8.9%; Score 89.8; DB 8; Length 757;
Best Local Similarity 50.6%; Pred. No. 9.3e-09;
Matches 217; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
121  GTGCAACAGCTGATTCGCGCGCTGCGGAGCTGGCGGACCTGGCCAAATGCGCGCGAG 180
Db      GCGGATGATCTCTGGGCTACGCGCTGATGTATGAGCAAGTACGCCAAGATCCAGTAAAC 571
Qy      CAGGTGATGCAATGCGCTCCGAAAGCATCACAAGAGACACTGCTCAAGCTGGGAAAC 240
Db      570  AACTTATCCAAAGCTTCATCGACTACATGAGGCGCTGAGCGCTGGCTCCAGCTCATGTCT 511
Qy      241  AGCGTGGCCGAGCTGGCCGACAGCAATGACGTGATGATCGTCAATCAACCAATGACCC 300
Db      510  GGTGTGCAAGAGCGCTATCAACAGTCAAGAGCGTTGCTGGGTGATCTTTCTCATGCAACA 451
Qy      301  GACACCTGTGAAGAAACCGCCTACTTTTGAACCTGTGGAAGAAAGCCGACCAAGCCGATC 360
Db      450  GATACGCTGAGAGAAACAGCTTTTGGCTTGATCTGACAGTGAAGTCCAAATTAACCAATC 391
Qy      361  GTTCGTGATGCTTCATGCGCCCGGACCGCCATGTCGCGCGAGCGCATGCTCAACTG 420
Db      390  GTTCGTGATGCTTCATGCGCGCAACGCGCTCGGTTTCAGACTTTGACGTCGCGGCAACTTG 331
Qy      421  TACAAGCGCGTGGCGGTGCGCAAGCAAGAGACTCGCGCGGCAAGGGCGGTGCTGTGACC 480

```

```

Db      330  CTTAAGCTGTGCGCATCGCTGTTGATCCCAATCCAAAGATCGTGGCCGATGCTGGCG 271
Qy      481  ATGAACGAGAGATTCAGTCCGGCGCTGACGTGAGAGAAATTCATCAACTCAAGCCGA 540
Db      270  ATGAACCAACCAATCAACGCTACCAAGTGTGTACACAGACACACTGCAACGTTGA 211
Qy      541  GCCTTCAG 549
Db      210  ACTTTTAC 202
RESULT 8
BH371112/c
LOCUS
DEFINITION   AG-ND-137P15.TR ND-TAM Anopheles gambiae genomic clone
ACCESSION   BH371112
VERSION      BH371112.1  GI:17317237
KEYWORDS
SOURCE
ORGANISM      Anopheles gambiae (African malaria mosquito)
               Anopheles gambiae
               Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
               Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
               Anopheles.
REFERENCE     1 (bases 1 to 663)
               Hong, Y.S., Hogan, J.R., Wang, X., Sarkar, A., Sim, C., Loftus, B.J.,
               Ren, C., Huff, E.R., Carlile, J.L., Black, K., Zhang, H.-B.,
               Gardner, M.J. and Collins, F.H.
               Construction of a BAC library and generation of BAC end
               sequence-tagged connectors for genome sequencing of the African
               malaria mosquito Anopheles gambiae
JOURNAL      Mol. Genet. Genomics 268 (6), 720-728 (2003)
MEDLINE      22542063
PUBMED       12655398
COMMENT       Other GSSs: AG-ND-137P15.TR
               Contact: Brendan J Loftus
               Department of Eukaryotic Genomics
               The Institute for Genomic Research
               9712 Medical Center Dr., Rockville, MD 20850, USA
               Tel: 301 838 0208
               Fax: 301 838 3543
               Email: b.loftus@igr.org
               This clone is from an A. gambiae BAC library (ND-TAM) provided by
               P.H. Collins and sequenced by The Institute for Genomic Research
               (TIGR). The BAC library was generated from A. gambiae PEST strain
               DNA. All DNA was extracted from newly hatched first instar larvae
               to minimize the inclusion of DNA from microorganisms that inhabit
               the gut. The DNA is derived from mixed sexes of larvae. The BAC
               library was constructed at Texas A&M University BAC Center
               University, College Station, Texas 77843-2123, USA using a HindIII
               partial digest.
               Seq primer: M13 For
               Class: BAC ends.
FEATURES
source       Location/Qualifiers
               1..663
               /organism="Anopheles gambiae"
               /mol_type="genomic DNA"
               /strain="PEST"
               /db_xref="taxon:7165"
               /clone="AG-ND-137P15"
               /clone_1lb="ND-TAM"
               /note="Vector: pECBAC1; Site_1: HindIII"
ORIGIN
Query Match      8.3%; Score 84.2; DB 8; Length 663;
Best Local Similarity 49.7%; Pred. No. 1.4e-07;
Matches 215; Conservative 0; Mismatches 218; Indels 0; Gaps 0;
134  TTGCGGCGTGGCCGAGCTGGCCGACCTGGCCAAATGTCGCGCGGACGAGCGTATGACGA 193
Db      618  TTGGCTACCGTGCCTGTGTAGGCAAGTACGCCAAGATCCAGGTAAACAATTATCCAAAG 559
Qy      194  TGGCTCCGAAAGCATCAACAGAGCAACTGTCTCAAGCTGGCAAGCAGCGTGGCCGAGC 253

```

Db 558 TTCCATCGGACTACATGAGGAGCGCTGGGTCAGCTCAGCTCTGCTGTGCAAGAGC 499
 Qy 254 TGGCCGACAGCAATACGTCGATGTCATGATCAACCCATGAGCAACCCGCAAG 313
 Db 498 CTATCAACAGTCAAGCTTGGCTGGTGATGTTCTCATGGCAAGATACGTCGAGG 439
 Qy 314 AAACCGCTACTTTTGAACCTGCTGGAAGAAAGACGACAGCCGATCGTGTGCTT 373
 Db 438 AAACAGCTTTTGGCTGTGATCTGACAGTGAATCCATTAACCATTCGTCGATTTGGG 379
 Qy 374 CCATGCGCCCGGCAACCCGATGTCGCGGAGGAGGCGTGTGTCGATGAACAGAGAG 433
 Db 378 CACAGCGCAACGCTCGCTTTCAGATTGACGCTCCGCAACTTGTCTTAACGCTGTGC 319
 Qy 434 CCGTGGCCAGCAACAGACTGCGCGGAGAGGCGTGTGTCGATGAACAGAGAGAG 493
 Db 318 GATGCTGTGTATCCCATTCAGAGATGTCGCGCCATGCTGGGATGAACAGAGAG 259
 Qy 494 TCCAGTCCGCGGCTGACGTCAGCAAGTCGATCAATCAACAGCCGATTCAGAGAGC 553
 Db 258 TCAAGCTACAGATGATGTCACCAAGACACACATGCCAAGCTTAACTTTAAGTTCT 199
 Qy 554 CCTGGGCGCCGCT 566
 Db 198 GGTGAGTTCGCT 186

RESULT 9
 CL694348 787 bp DNA linear GSS 10-UTL-2004
 LOCUS PRI0164a A10.2 - PRI0164a.BR (787) Mixed stage fosmid library of P.
 DEFINITION pacificus var. California Pristionchus pacificus genomic, genomic
 survey sequence.

ACCESSION CL694348
 VERSION CL694348.1 GI:50216256
 KEYWORDS Pristionchus pacificus
 SOURCE Pristionchus pacificus
 ORGANISM Pristionchus pacificus

REFERENCE 1 (bases 1 to 787)
 AUTHORS Strimvasan,J., Otto,G.W., Kahlow,U., Gelsler,R. and Sommer,R.J.
 TITLE AppDB: an AcedB database for the nematode satellite organism
 Pristionchus pacificus

JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)
 COMMENT Contact: Sommer RJ
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498

EMAIL: ralf.sommer@uebingen.mpg.de
 This library was generated at Caltech, Pasadena, USA and end
 sequenced at Vancouver, Canada.
 Seq primer: T7
 Class: fosmid ends.

FEATURES
 source Location/Qualifiers

1..787
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus
 var. California"
 /note="Vector: pBplfos-5 Fosmid vector"

ORIGIN

Query Match 8.3%; Score 84; DB 9; Length 787;
 Best Local Similarity 49.0%; Pred. No. 1.6e-07;
 Matches 253; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

502 GGGCGTACGTGACAGTCAATCAATCAAGACCAAGCCTTCAAGAGCG---CCTGG 558

Db 778 GGGCGTATGTCACCAAAACCAACACCGCCATGTAGGACCTTCAATCTGTATCAAC 719
 Qy 559 GGGCCGCTGGGATGTGTGTAAGGCAAGTCGATCTGTTCCGCTCCGCGCAAGCGC 618
 Db 718 GATCTCTGGGTTACATTCACACGGTGAATGATCACTCAACAGTACCCCGCAGTAG 659
 Qy 619 CACAGGTCATCTCGAGTTGACATCAAGCAATCAGACGCTGCCAAGGTGACATC 678
 Db 658 CACACCAAGCAACAGCGCTGATGATCTTAAGCTGAATGAATGACCCGAAAGTCGCAAT 599
 Qy 679 GCCTACAGTATGACAGTCAACGACGCGCTTACAGAGCCCTTGGCACAAGACGGCGCC 738
 Db 598 GTTATATACAGCTTAAGCAATCCGATCTTCGCTTAAGCACTGTGATGATCGGCTAT 539
 Qy 729 AAGGCGCTGATCAATGCGGACCGCAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
 Db 538 GATGCACTGTTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
 Qy 739 CTGACAGAGCTGCGCAGAACGCGCTGCAATCATTCGTTGCTCACTCAACAGGCGGT 858
 Db 478 CTGCAACCGCTGCGCAAAACGCGCACTGCAATGATGCTTCCGCGTACCGAGCGGC 419
 Qy 859 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 918
 Db 418 GCTACCACTCAGAGATGCGCAAGTATGATGCAATGATGATGATGATGATGATGATGAT 359
 Qy 919 CTGAACCGGCAAGAGCGCGCATCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978
 Db 358 TTGAACCGGCAAAAGCGCGCTTCTGCTGCACTGCTGCTGCAACCACTTAAGATCGG 299
 Qy 979 AAGAGCTGACAGCGCATTTTCTGGAATACTGATTA 1014
 Db 298 CAGCAGATCCAGCAGATCTTCATCATCACTAATCA 263

RESULT 10
 AQ90947 724 bp DNA linear GSS 14-AUG-2000
 LOCUS Rfco1794 Photorhabdus luminescens strain W14 M13 library
 DEFINITION Photorhabdus luminescens genomic clone PLG01794, genomic survey
 sequence.

ACCESSION AQ90947
 VERSION AQ90947.1 GI:9649541
 KEYWORDS Photorhabdus luminescens
 SOURCE Photorhabdus luminescens
 ORGANISM Photorhabdus luminescens

REFERENCE 1 (bases 1 to 724)
 AUTHORS ffrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,
 Daborn,P.J., Bowen,D. and Blattner,F.R.
 TITLE A genomic sample sequence of the entomopathogenic bacterium
 Photorhabdus luminescens W14: potential implications for virulence
 Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)

JOURNAL MEDLINE
 PUBMED
 COMMENT Contact: ffrench-Constant RH
 Department of Biology and Biochemistry
 University of Bath
 South Building, Bath BA2 7AY, UK
 Tel: (44) 1225 826621
 Fax: (44) 1225 826779
 Email: bsarfbath.ac.uk
 This is one of 2,122 random reads from the M13 library. For
 annotation of identified clones (BLASTX, BLASTN and mapping to E.
 coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic
 Acids Res.

Seq primer: M13 Forward
 Class: shotgun.

FEATURES
 source Location/Qualifiers

1..724
 /organism="Photorhabdus luminescens"

/mol_type="genomic DNA"
/strain="W14"
/db_xref="taxon:29488"
/clone="pUG01794"
/dev_stage="primary phase variant"
/clone_lib="Photocorhabdus luminescens strain W14 M13 library"
/note="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janss."

ORIGIN

Query Match 8.1%; Score 82; DB 8; Length 724;
Best Local Similarity 55.1%; Pred. No. 4.2e-07;
Matches 207; Conservative 0; Mismatches 160; Indels 9; Gaps 2;

QY 44 TGGCCACCGGCGGACCATGCGGCGCTGGCGCCAGCGGCGCAACGCGCACTACC 103
DB 379 TGGCGACGGGGGAGCTATGCGGG--TGGCGGAAATCTGCATTCATTCATTTATG 323
QY 104 AGGCTGCCAAGTTGGGTGCAACAAGCTATTGCCGCGTGGCGGAGCTGGCGGACTGG 163
DB 322 TTCTGTGAAGGGTTGCTATTGAAATTAATTAATTAATTAATTAATTAATTAATTA 263
QY 164 CCAATGTGCGCGCGAGCAGTGTATGCAATCGCTCCGAAAGCATCAACAGAGACC 223
DB 262 CCAACTAAAGGTGAGCAGGTAGTTAGTTAGTTGTTCCGAGATATGATCAAGTGT 203
QY 224 TGTCTCAAGTGGCAAGCAGGTGGCGGAGCTGGCGGCAAGATGACGTGATGCAATG 283
DB 202 GGTGTGACGTGAGGAAGAAATCAAC-----GCTGATGTGCAAAACCGATGTTTG 149
QY 284 TCATCAACCAATGACCGACACCGCTGGAAGAAACCGCTACTTTTGAACCTGTGGA 343
DB 148 TTTATCCCAAGGTATCTGACACATGAGAAACCGTTTCTTCTGATCTCACTAAC 89
QY 344 AGACCGCAAGCCGATGTCGTGTCGTTTCATGCGCCCGGCGGCGGCGGCTGCGCG 403
DB 88 ACTGCCAAGACCGGATGATGATGTTGGCGGCAATGCGGCACTAAGCGCTTGTGCG 29
QY 404 ACCGCAATGCTCAACT 419
DB 28 ATGGCCCGTTAAATCT 13

RESULT 11
AQ500605/c 766 bp DNA linear GSS 29-APR-1999
LOCUS V36C6 mtn-3xHA/lacZ insertion library Saccharomyces cerevisiae
DEFINITION genomic 5', genomic survey sequence.
ACCESSION AQ500605
VERSION AQ500605.1 GI:4706315
KEYWORDS GSS.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetiales; Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 766)
Rosa-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desBergues, S.A., Cheung, K.-H., Sheehan, A., Symoniatidis, D., Jansen, R.,
Umanaky, L., Heldman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
Unpublished (1999)
CONTACT: Kumar A
JOURNAL: Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumaryale.yale.edu
te of mtn-3xHA/lacZ insertion.
Seg primer: GGCCTTCTTCTTGGAAATAC

Class: transposon-tagged.
Location/Qualifiers
FEATURES
source
1..766
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/lab_host="S. coli"
/clone_lib="mtm-3xHA/lacZ insertion library"
/note="Vector: pHS6-Sal; A yeast genomic DNA library
(lacking mitochondrial DNA) was prepared in pHS6-Sal;
genomic DNA was size-fractionated (DNA of roughly 2-3 kb
in length) prior to cloning. This library was
subsequently mutagenized with a mtn-3xHA/lacZ
mitransposon containing lacZ, URA3, and tet resistance."

ORIGIN

Query Match 8.0%; Score 81; DB 8; Length 766;
Best Local Similarity 50.6%; Pred. No. 6.9e-07;
Matches 195; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 120 CCGTCAAGAGTATGCGGCGTGGCGGAGCTGGCGGACCTGCGCAATGTGGCGGCA 179
DB 534 CGTATATGATTTAAATGAAGCGCTCCATCTTAAGCTGGAAGGAGATCTGATATCT 475
QY 180 GCAGTATGACAGATGCGCTCCGAAGATCAACAGAGACCTGCTCAAGCTGGCAAG 239
DB 474 TCAAGTGTCAAGTGTGTTCACTTTTAACTTAAGCATCTATCATTTGATATCA 415
QY 240 CAGCGTGGCGAGCTGGCGGAGCAAGCAATGACGTGATGCGATGTCATGACCCAGC 299
DB 414 CGGTATCTCGAGGCACTAGCTCTGATGATTAAGCTGAGTGGGTGTCATCATGGGAC 355
QY 300 CCAACACCTGGAAGAAACCGCTTATTTTGAACCTGTGGAAGAAACCGCAAGCCGAT 359
DB 354 CCACTATGAGGAGAAACAGCTTTCTTGAATTTGACATTAATTCAGAAAGCAGT 295
QY 360 CGTGTGTGTCGTGTCATGCGCGCGGCAACCGCATGTCCGCGGAGGAGTCTCAACT 419
DB 294 ATGTATCGAGGCGCTATGCTGTCAGCCACTGCGCATGTGCTGATGGCCCATGATTT 225
QY 420 GTTCAAGCGCGTGGCGGCGGCAAGCAAGACTGCGCGGAGGCGGTGCTGATAC 479
DB 234 ATATCAAGCAGGTGTATGCTGCTTCTGGAATCACTGGGTCGATGATAC 175
QY 480 CATGAACGACGAGATCAAGTCCGGG 504
DB 174 TCTAAACGATCGTATGCTCTGGG 150

RESULT 12
AQ500657/c 843 bp DNA linear GSS 29-APR-1999
LOCUS V34F11 mtn-3xHA/lacZ insertion library Saccharomyces cerevisiae
DEFINITION genomic 5', genomic survey sequence.
ACCESSION AQ500657
VERSION AQ500657.1 GI:4706367
KEYWORDS GSS.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetiales; Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 843)
Rosa-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desBergues, S.A., Cheung, K.-H., Sheehan, A., Symoniatidis, D., Jansen, R.,
Umanaky, L., Heldman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
Unpublished (1999)
CONTACT: Kumar A
JOURNAL: Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA

Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mtn-3xHA/lacZ insertion.
 Seg primer: GGCTCTTCTTGGAAAGTAC
 Class: transposon-tagged.
 Location/Qualifiers

FEATURES

source

1. 843
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_1lb="mtn-3xHA/lacZ insertion library"
 /note="Vector: pHS86-Sal; A yeast genomic DNA library
 (lacking mitochondrial DNA) was prepared in pHS86-Sal;
 genomic DNA was size-fractionated (DNA of roughly 2-3 kb
 in length) prior to cloning. This library was
 subsequently mutagenized with a mtn-3xHA/lacZ
 minitransposon containing lacZ, URA3, and tetr resistance."

ORIGIN

Query Match 7.8%; Score 78.6; DB 8; Length 843;
 Best Local Similarity 49.9%; Pred. No. 2.3e-06;
 Matches 192; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

```

QY 120 CCGTCAAGCTGATTCGCGCGGAGCTGCGGACCTGCGCAATGCGCGCGA 179
DB 671 CCGTAATGATTTATATAGAGCCCTCCATCTTATGCTGAGAGGACATCTGACATATCT 612
QY 180 GCGAGTGATGAGATGCGCTCCGAAAGCATCAACGACGACCTGCTCAAGTGGCAAG 239
DB 611 TCAAGTGCTAAAGTGGTTCAATCTTATTAATGACGCACTGATCCCATTTGATCA 552
QY 240 CAGCGTGGCGGAGCTGCGGACGACGATGATGATGATGATGATGATGATGATGATGAT 299
DB 551 CCGTATCTCCGAGGACCTAGCTTCTGATGATGATGATGATGATGATGATGATGATGAT 492
QY 300 CGACACCTGGAAGAAACCGCTACTTTTGAACCTGCTGGAAGAAAGACGACGACGAT 359
DB 491 CGACATATGAGAGAGACAGCTTTCTTTAGATTGACATTAATTTGAGAGAGCAAT 432
QY 360 CCGTGGTGGTTCATGCGCGCGCGGACGACGATGCGCGGACGACGATGCTCAACT 419
DB 431 ATGTATGCGAGCGGCTATGCGTCAAGCACTGCGCACTGATGCGCGCAATGATTT 372
QY 420 GTTCAAGCGCGTTCATGCGCGCGGACGACGATGCGCGGACGACGATGCTCAACT 479
DB 371 ATATCAAGCGGCTATGCGTCTTCTGAGAAATCACTGCGGCTGCGGACGATGATCA 312
QY 480 CATGAACGACGAGATCCAGTCCGGG 504
DB 311 TCTAAACGATCGATTCCTCTGGG 287

```

RESULT 13
 AOS01722/c 453 bp DNA linear GSS 29-APR-1999
 LOCUS V15G1 mtn-3xHA/lacZ Insertion library Saccharomyces cerevisiae
 DEFINITION genomic 5', genomic survey sequence.
 ACCESSION AOS01722
 VERSION AOS01722.1 GI:4707372
 KEYWORDS GSS.
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 1 (bases 1 to 453)
 Roser-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
 deBragas, S.A., Cheung, K.-H., Sheehan, A., Symonakis, D., Jansen, R.,
 Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
 Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
 Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
 Gene Disruption

JOURNAL

Unpublished (1999)
 Contact: Kumar A

COMMENT

Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mtn-3xHA/lacZ insertion.
 Seg primer: GGCTCTTCTTGGAAAGTAC
 Class: transposon-tagged.
 Location/Qualifiers

FEATURES

source

1. 453
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_1lb="mtn-3xHA/lacZ insertion library"
 /note="Vector: pHS86-Sal; A yeast genomic DNA library
 (lacking mitochondrial DNA) was prepared in pHS86-Sal;
 genomic DNA was size-fractionated (DNA of roughly 2-3 kb
 in length) prior to cloning. This library was
 subsequently mutagenized with a mtn-3xHA/lacZ
 minitransposon containing lacZ, URA3, and tetr resistance."

ORIGIN

Query Match 7.3%; Score 73.6; DB 8; Length 453;
 Best Local Similarity 52.6%; Pred. No. 2.5e-05;
 Matches 160; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

```

QY 248 CCGAGCTGGCCGACGACATGATGATGATGATGATGATGATGATGATGATGATGAT 307
DB 366 CCGAGGACCTAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 307
QY 308 TGGAGAAACCGCCCTACTTTTGAACCTGCTGGAAGAAAGACGACGACGATGCTG 367
DB 306 TGGAGAGACAGCTTCTTCTTGAATTTGACATTAATTTGAGAGAGCAATGATGATG 247
QY 368 TCGGTTCCATGCGCGCGGACGCGCATGTCGCGGACGCGATGCTCAACTGTAACAG 427
DB 246 CAGCGGCTATGCTGCTCAAGCACTGCGCACTGCGGATGCGGACGATGATTTATCAAG 187
QY 428 CCGTGGCCGTGGCCAGCAAGAACTGCGCGGCGGAGGCGCTGCTGATGACATGAGC 487
DB 186 CAGTGTCTATGCTCTCTCTGGAATCACTGCGGCTGCGGACGATGACATCTAAAG 127
QY 488 ACGAGATCCAGTCCGCGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCG 547
DB 126 ATGTAATGCTCTGCGGTTTGGACACGAAATGAAATGCCAATCTTTAGATACATTA 67
QY 548 AGAG 551
DB 66 GGGG 63

```

RESULT 14
 AOS01723/c 459 bp DNA linear GSS 29-APR-1999
 LOCUS V15G2 mtn-3xHA/lacZ Insertion library Saccharomyces cerevisiae
 DEFINITION genomic 5', genomic survey sequence.
 ACCESSION AOS01723
 VERSION AOS01723.1 GI:4707373
 KEYWORDS GSS.
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 1 (bases 1 to 459)
 Roser-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
 deBragas, S.A., Cheung, K.-H., Sheehan, A., Symonakis, D., Jansen, R.,
 Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
 Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
 Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
 Gene Disruption

REFERENCE
 AUTHORS
 TITLE

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 21:13:22 ; Search time 168 Seconds
(without alignments)
773.521 Million cell updates/sec

Title: US-09-842-628-2

Perfect score: 1680

Sequence: 1 KEVENQQLANVILATGST.....MVAMTKQDSKELQRIFMWEY 336

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp19808:*\n2: geneseqp19908:*\n3: geneseqp20008:*\n4: geneseqp20018:*\n5: geneseqp20028:*\n6: geneseqp20038:*\n7: geneseqp20038:*\n8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1680	100.0	336	2	AAR59739
2	1432.5	85.3	393	7	ABO73618
3	1063.5	63.3	374	6	ADA35175
4	782	46.5	348	3	AAW90867
5	758	45.1	375	8	AD136574
6	749	44.6	327	2	AAR21555
7	749	44.6	348	1	AA703348
8	749	44.6	348	1	ABP98617
9	717	42.7	364	2	AAW98609
10	711	42.3	356	7	ABO61384
11	710.5	42.3	346	7	ADP05106
12	685.5	40.8	348	6	ABM68938
13	677.5	40.3	326	6	ABP98616
14	677.5	40.3	348	6	ABP98615
15	602.5	35.9	378	8	ADDO07179
16	560.5	33.4	378	8	ADDO07181
17	559	33.3	374	8	ADDO07183
18	536.5	31.9	378	8	ADN97130
19	535.5	31.9	451	8	ADDO07187
20	528	31.4	371	8	ADDO07185
21	517	30.8	379	8	ADDO07189
22	485	28.9	362	8	ADDO07190
23	465	27.7	100	2	AAK37660
24	372.5	22.2	333	7	ADH86728
25	361.5	21.5	322	4	AAK82409

26	361.5	21.5	326	5	ABP38389	ABP38389 Staphyloc
27	360.5	21.5	323	5	ABBS4049	ABBS4049 Lactococc
28	356	21.2	320	6	ABU02502	ABU02502 S. pneumo
29	356	21.2	320	8	ADK47071	ADK47071 Streptoco
30	356	21.2	320	8	ADM92242	ADM92242 S. pneumo
31	356	21.2	321	8	ADR94759	ADR94759 Novel S.
32	349.5	20.8	322	6	ABM70762	ABM70762 Staphyloc
33	332	19.8	321	5	ABP27237	ABP27237 Streptoco
34	331.5	19.0	324	7	ADC95330	ADC95330 R. faeciu
35	318.5	19.0	330	5	ABP27236	ABP27236 Streptoco
36	317	18.9	320	5	ABBS9122	ABBS9122 Listeria
37	300	17.9	297	5	ABP30464	ABP30464 Streptoco
38	294.5	17.5	114	5	ABU51451	ABU51451 Helicobac
39	291	17.3	450	8	ADN46622	ADN46622 Thermococ
40	279	16.6	439	4	AAK96655	AAK96655 Putative
41	269.5	16.0	458	7	ADM26903	ADM26903 Hytherther
42	251	14.9	325	4	AAK79696	AAK79696 Corynebact
43	251	14.9	325	4	AAK92092	AAK92092 C. glutam
44	226.5	13.5	340	4	AAK96780	AAK96780 Putative
45	216.5	12.9	328	8	ADN46369	ADN46369 Thermococ

ALIGNMENTS

RESULT 1

AAR59739 standard; protein; 336 AA.

AC	XX	AAR59739;	
AC	XX		
DT	XX	25-MAR-2003 (revised)	
DT	XX	12-JUN-1995 (first entry)	
DE	XX	Pseudomonas glutaminase.	
DE	XX		
XX	XX	Glutaminase; antiviral; virucide; anticancer; cancer therapy; HIV virus;	
KM	XX	Gene therapy; Escherichia coli.	
XX	XX		
OS	XX	Pseudomonas sp.	
XX	XX		
PN	XX	W09413817-A1.	
XX	XX		
PD	XX	23-JUN-1994.	
XX	XX		
PF	XX	04-DEC-1992; 92WC-US010421.	
XX	XX		
PR	XX	04-DEC-1992; 92WC-US010421.	
XX	XX		
PA	XX	(MEME-) MB MEDICAL ENZYMES AG.	
XX	XX		
PI	XX	Roberts J, Macallister TW, Sethuraman N, Freeman AG;	
XX	XX		
DR	XX	WPI; 1994-217891/26.	
DR	XX	N-PSDB; AAK68438.	
XX	XX		
PT	XX	Recombinant glutaminase derived from Pseudomonas 7A - expressed in E.	
PT	XX	coli to increase yield and avoid Pseudomonas endotoxins for antiviral and	
XX	XX	anticancer therapy.	
XX	XX		
PS	XX	Disclosure; Page 33-34; 60pp; English.	
XX	XX		
CC	XX	Chromosomal DNA from Pseudomonas sp. 7A (ATCC 29598) was used to	
CC	XX	construct a genomic library in Escherichia coli LB392. Screening with	
CC	XX	mixed oligonucleotide probes was used to isolate a glutaminase- encoding	
CC	XX	clone. This was sequenced using the primers given in AAK68439-47. The	
CC	XX	gene can be used to manufacture recombinant glutaminase, free of	
CC	XX	Pseudomonas exotoxin, for use in e.g. HIV and cancer therapy. The gene	
CC	XX	may also be used in gene therapy protocols. (Updated on 25-Mar-2003 to	
CC	XX	correct PN field.)	
XX	XX	Sequence 336 AA;	
SQ	XX		

Query Match 100.0%; Score 1680; DB 2; Length 336;
Best Local Similarity 100.0%; Pred. No. 5.7e-131;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEVENOQKLANVILATGTTAGAGASANSATYQAAGVYKLAGVPELADLANVGE 60
DB 1 KEVENOQKLANVILATGTTAGAGASANSATYQAAGVYKLAGVPELADLANVGE 60
QY 61 QVQJASESITNDLLKLASSVAELADSNVDGIVITHTGDTLBEETAYFLNVEKTDKPI 120
DB 61 QVQJASESITNDLLKLASSVAELADSNVDGIVITHTGDTLBEETAYFLNVEKTDKPI 120
QY 121 VVVGSMRPGTAMSDGMLNLNAVAVASNDSRGKGVLTVMNDEIQSGRDVSKSINIKTE 180
DB 121 VVVGSMRPGTAMSDGMLNLNAVAVASNDSRGKGVLTVMNDEIQSGRDVSKSINIKTE 180
QY 181 AFKSMGPIGMYVEKSYWFRLPARKHTVNSEFDIKQISSLPQVDIAYSYGAVTDTAYKA 240
DB 181 AFKSMGPIGMYVEKSYWFRLPARKHTVNSEFDIKQISSLPQVDIAYSYGAVTDTAYKA 240
QY 241 LAONGAKALIHAGTNGSVSSRVVPALQELRNQVQIIRSSRQGGFVLRNAEQPDNDK 300
DB 241 LAONGAKALIHAGTNGSVSSRVVPALQELRNQVQIIRSSRQGGFVLRNAEQPDNDK 300
QY 301 WVVAHDLPQKARILAMVAMTKTQDSKEIQRIFWEX 336
DB 301 WVVAHDLPQKARILAMVAMTKTQDSKEIQRIFWEX 336

RESULT 2

ABO73618 standard; protein; 393 AA.

ABO73618;

29-JUL-2004 (first entry)

Pseudomonas aeruginosa polypeptide #5793.

Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

Pseudomonas aeruginosa.

US6551795-B1.

22-APR-2003.

18-FEB-1999; 99US-00252991.

18-FEB-1998; 98US-0074788P.

27-JUL-1998; 98US-0094190P.

(GENO-) GENOME THERAPEUTICS CORP.

Rubenfeld MJ, Nolling J, DeLoughery C, Bush D;

WPI; 2003-615309/58.

N-PsDB; ABD07189.

Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

Dielschauer; SEQ ID NO 22364; 455pp; English.

The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant

CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

Sequence 393 AA;

Query Match 85.3%; Score 1432.5; DB 7; Length 393;
Best Local Similarity 84.3%; Pred. No. 2.6e-110;
Matches 284; Conservative 22; Mismatches 30; Indels 1; Gaps 1;

QY 1 KEVENOQKLANVILATGTTAGAGASANSATYQAAGVYKLAGVPELADLANVGE 60
DB 57 KEVENOQKLANVILATGTTAGAGASANSATYQAAGVYKLAGVPELADLANVGE 116
QY 61 QVQJASESITNDLLKLASSVAELADSNVDGIVITHTGDTLBEETAYFLNVEKTDKPI 120
DB 117 QVQJASESITNDLLKLASSVAELADSNVDGIVITHTGDTLBEETAYFLNVEKTDKPI 176
QY 121 VVVGSMRPGTAMSDGMLNLNAVAVASNDSRGKGVLTVMNDEIQSGRDVSKSINIKTE 180
DB 121 VVVGSMRPGTAMSDGMLNLNAVAVASNDSRGKGVLTVMNDEIQSGRDVSKSINIKTE 236
QY 177 VVVGSMRPGTAMSDGMLNLNAVAVASNDSRGKGVLTVMNDEIQSGRDVSKSINIKTE 240
DB 181 AFKSMGPIGMYVEKSYWFRLPARKHTVNSEFDIKQISSLPQVDIAYSYGAVTDTAYKA 296
QY 237 AFKSMGPIGMYVEKSYWFRLPARKHTVNSEFDIKQISSLPQVDIAYSYGAVTDTAYKA 296
DB 241 LAONGAKALIHAGTNGSVSSRVVPALQELRNQVQIIRSSRQGGFVLRNAEQPDNDK 299
QY 297 LAONGAKALIHAGTNGSVSSRVVPALQELRNQVQIIRSSRQGGFVLRNAEQPDNDK 356
DB 300 DMVVAHDLPQKARILAMVAMTKTQDSKEIQRIFWEX 336
QY 357 DMVVAHDLPQKARILAMVAMTKTQDSKEIQRIFWEX 393

RESULT 3

ADA35175 standard; protein; 374 AA.

ADA35175;

20-NOV-2003 (first entry)

Acinetobacter baumannii protein #2336.

Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;

plant biocontrol agent.

Acinetobacter baumannii.

US6562958-B1.

13-MAY-2003.

04-JUN-1999; 99US-00328352.

09-JUN-1998; 98US-0088701P.

(GENO-) GENOME THERAPEUTICS CORP.

Breton G, Bush D;

WPI; 2003-576092/54.

N-PsDB; ADA31049.

New Acinetobacter baumannii proteins and nucleic acids, useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, or as biocontrol agents for

plants.

Example, SEQ ID NO 6462; 328bp; English.

The invention relates to isolated *Acinetobacter baumannii* nucleic acids. The *A. baumannii* nucleic acids and polypeptides are useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, to detect the presence of *A. baumannii* and other *Acinetobacter* species in a sample, in screening compounds for the ability to interfere with the *A. baumannii* life cycle or to inhibit *A. baumannii* infection, and as biocontrol agents for plants. The present sequence represents the amino acid sequence of an *A. baumannii* protein.

Sequence 374 AA;

Query Match 63.3%; Score 1063.5; DB 6; Length 374;
Best Local Similarity 64.7%; Pred. No. 1.1e-79;
Matches 213; Conservative 45; Mismatches 68; Indels 3; Gaps 2

11 NVVLTATGTTAGAGASAANSATYQAAKVGVDKLTAGVPELADLANVGEQVMQIASBSI 70
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
46 NVVVATGTTAGAGASSANSATYPAKPPVALINAVFOIDLNAVSGIQALQVASSBSI 105
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
71 TNDLLKLAASVAEELADSDVDGVIYTTGCTDTLBEFAVFLINVEKTDKPIVVGSMRPST 130
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
106 TDKEILQIARQVNEIWKKEPTVAGVVIYTHGDTLBEFAPEFLINVHTDKPIVVGSMRPST 165
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
131 AMSADGMINTLYNAVAVASNKDSRGKGLVTWMDRIQSGRDVSKSINIKTEAFKSAWGPIG 190
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
166 ALSADGPINLVSAVALAASDDAKKNGWYLMWDSIFPAARDYKGINIHNAFVSQMGALG 225
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
191 NVVEGKSYWFLPAKRHTVNSSEFDIKQI--SSLPQVDIAVSYGNVTDRAYYKLAONGAKA 248
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
226 TLVEGKPYWFRQSVKHTVASEFENIKGDLPTQIYVYGSDSMLPDAYEAYAKAKAKA 285
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
249 LIHAGTNGSVSSKRVVPALQELR-KNGVYIINSSRQGGPFLRANRQPPDKNDVYVAHL 307
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
286 IHHAGTNGSVAKYIVPTLQNIHDKNGIGIIRSSRVQGFVLRDARQPSKYGWAAHDL 345
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
308 NPQKRIILAMVAMTKTKDSKELQRIFFMEY 336
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
346 NPQKRLIALALTLTINDAKGIQRMFWQY 374
:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

RESULT 4
AAM90867
AAM90867 standard; protein, 348 AA.
XX AAM90867;
XX AC
XX 30-JUN-2000 (first entry)
XX DT
XX L-asparaginase protein.
XX DB
XX L-asparaginase; fermentation; ss.
XX KW
XX L-asparaginase; fermentation; ss.
XX OS
XX unidentified.
XX PN
XX CN1237633-A.
XX PD
XX 08-DEC-1999.
XX PE
XX 01-JUN-1998; 98CN-00102046.
XX PR
XX 01-JUN-1998; 98CN-00102046.
XX PA
XX (MCCR-) INST MICROBIOLOGY CHINESE ACAD SCI.
XX PI
XX Qian S, Wang Y, Meng G;
XX WPI, 2000-351194/31.
XX N-PSDB; AAX82834.
XX DR

XX	New structured L-asparaginase bacterial host cell and its production culture.
PT	
PS	Disclosure; Fig 2; 12pp; Chinese.
CC	This invention describes a novel L-asparaginase-producing recombinant bacteria and its fermentation culture method. The new cell is produced using standard recombinant techniques. The enzyme gene expressed in the cell is different from the reported enzyme gene in sequence, and its individual amino acid composition is also different. The culture medium available for culturing the new bacterial cells is cheap and easily express the enzyme and the fermentation process used to grow the cells and represents an L-asparaginase which is described in the method of the invention
CC	
CC	
SQ	Sequence 348 AA;
Query Match	46.5%; Score 792; DB 3; Length 348;
Best Local Similarity	48.3%; Pred. No. 2,3e-56;
Matches 159; Conservative	59; Mismatches 107; Indels 4; Gaps 3;
Df	
Qy	9 LANVTIATGGTIGAGASANSAYQAQKGVNKLINGVELADLANVREGQVMQLASR 68 : : : : 23 LPNTITLTTGTTIGGGDSATKS -NTYRKGVENLVANAVQLKDIAVKGEQVNINISQ 81 : 69 SITNDLLKLASSVAELADSNVDGIVITHTGDTLEETAYFLNVKTDKPIVVGSNRP 128 : 82 DMNDVMTLTAKKIN--TDCKDQDFVITHGTDMETAYFLDLTVKCDKRVVMVGAMRP 139 : 129 GTMSADGMLNNAVANASKDRGKVLYTMDEIOSGDVSKSINIKTEAFRSA-WG 187 : 140 STMSADDPFVLNNAVVAADKASNRGLVVMNDVLDGHDVTKINTDVATFRSVNYG 199 188 PLGMWVEKSYWPLPAKHHTVNSBFDIKOISLPDYDAVSXGNVTEDTAVALAQNCAK 247 : : : : : : : : : : : : : : : : : : : : : : : : : : 200 PLSYTHNKIDYORTPAKHTSDPEPDYSKNELPKGIVTYNVANASDLPKALVDAGYD 259 248 ALIHAGTNGSVSSRVVPALOELEKNGVOILRSSROOGGFVLRANAEPDDKNDWVAHDL 307 : : : : : : : : : : : : : : : : : : : : : : : : Db 260 GIVSGVGNGLNLYSVPEPTLATATAKNGAVVRSSRPVPGATTQDAEVDADKYGFVASGL 319 308 NPQRRIIANVAMTKTODSKELORIFWEY 336 Qy 320 NPOKARVILLALQTOTKDPQOIIOQIFNQY 348 Db. RESULT 5 ADI36574 ID ADI36574 standard; protein; 375 AA. AC ADI36574; XX XX 15-APR-2004 (first entry) DX DE Bacillus subtilis L-asparaginase protein. XX KM Secreted protein; L-asparaginase; acute lymphocytic leukaemia; therapy; enzyme. XX OS Bacillus subtilis. XX FH Key Location/Qualifiers FH Peptide 1..23 FT /label= Signal_peptide FT Protein 24..375 PN /note= "Mature Bacillus subtilis L-asparaginase protein" XX FD 02-OCT-2003.

PF 01-APR-2003; 2003US-00406025.
 XX
 PR 01-APR-2002; 2002US-0369192P.
 XX
 XX (NOVO) NOVOZYMES BIOTECH INC.
 PA Thomas MD, Sloma A;
 XX
 PI
 XX WPI; 2004-088916/09.
 DR N-PSDB; ADI36573.
 XX
 PT Production of a secreted polypeptide having L-asparaginase activity for
 PT treating leukemia, comprises cultivating a host cell comprising a nucleic
 PT acid having a sequence encoding a secretory signal peptide linked to a
 PT second sequence.
 XX
 PS Claim 34; SEQ ID NO 2; 22pp; English.
 XX
 CC The present invention relates to recombinant methods for producing
 CC secreted polypeptides having L-asparaginase activity. The invention is
 CC useful for production of a secreted protein having L-asparaginase
 CC activity, for use in producing L-asparagine from L-asparagine. The
 CC invention is useful for treatment of leukemia such as acute lymphocytic
 CC leukemia. The present sequence is the Bacillus subtilis L-asparaginase
 CC protein.
 CC
 SQ Sequence 375 AA;
 Query Match 45.1%; Score 758; DB 8; Length 375;
 Best Local Similarity 47.6%; Pred. No. 2.5e-54;
 Matches 158; Conservative 58; Mismatches 112; Indels 4; Gaps 3;
 QY 6 QOKLNVLLATGTTAGAGASANSATYQAAKVGVDKLIAGVPELADLANVREGVMOI 65
 DB 47 KDLPIRILATGTTAGADQSTETKAGVGVESLIBAVPEKDIANVSGEIVNV 106
 QY 66 ASESITNDLLKLASSVAELADSNVDGIVITHGTDLSEETAFPLNVEKTPVIVVGS 125
 DB 107 GSTINDKILKLAKIRINHLASDDVDGIVTHGTDLSEETAFPLNVEKTPVIVVGS 166
 QY 126 KSPGTMSADGMLNVAVAVASNKDSRGKGLVTMNDIEQSDVSKSINIKTEAFKS- 184
 DB 167 KSPSTHLSADGFSNLNVAVAVAGAPKAGKGTLLVNDRIASARYITKTNTTIDFKSE 226
 QY 185 AWGPIGVVSGSYWFRLLPAKHTVNSSEPDIKOISLPQVDIAYSGNVTDTAYKALQN 244
 DB 227 EMGPGFTIAD-DIYFNNETIRKHTKOTDPSVSNLDELPOVDIIYQNDGSLFLPDAVKA 285
 QY 245 GAKALIHAGTNGSVSSRVVPLQELRKNGVQIIRSSRQGGFVLRNAEQPDDKNDVVYA 304
 DB 286 GAKGIYFAGSGSGSLSDAAKGAADSAVKKGVTVVSTRGNGVVTPEMDYAR--KDLAS 343
 QY 305 HDLNPOKARILLAVAMTKTQDSKELORIEMV 336
 DB 344 NSLNPOKARILLALTKTNDPOKIOAYFNEY 375
 RESULT 6
 AAR21555
 ID AAR21555 standard; protein, 327 AA.
 AC AAR21555;
 XX
 XX 10-JUN-1992 (first entry)
 DT
 XX
 XX L-asparaginase.
 KW Purification process; super-expression; lymphoblastic leukaemia.
 XX
 OS Erwinia chrysanthemi.
 XX
 XX WO9202616-A.
 PN
 XX

PD 20-FEB-1992.
 XX
 XX
 PF 02-AUG-1990; 90GB-00017002.
 XX
 PR 02-AUG-1990; 90GB-00017002.
 XX
 XX (PUBL-) PUBLIC HEALTH LAB S.
 PA
 XX
 PI Goward C;
 XX
 XX WPI; 1992-080071/10.
 DR
 XX
 XX Purificn. of Erwinia L-asparaginase - by contact with a solid medium
 PT having cation exchange groups, pref. sulphamate.
 PT
 XX
 PS Disclosure; Fig 1; 17pp; English.
 XX
 CC The protein sequence was obt. from a 12 hr. seed culture of E.
 CC chrysanthemi added to a medium contg. yeast extract and Na glutamate.
 CC After centrifugation the pH was adjusted and the supernatant further
 CC centrifuged to remove any deposit. The extract was purified by a S-
 CC Sepharose Fast Flow column, having cation exchange gpe. (comprising
 CC sulphamate gpe.) to absorb asparaginase on the support. A degree of
 CC purificn. resulting in a specific activity >600U/mg was achieved after 6
 CC purificn. steps with a yield of 40 percent. L-asparaginase can be
 CC produced of a very high quality from an amplified process which allows
 CC for increased automation. E. chrysanthemi is known to be effective
 CC against acute lymphoblastic leukaemia
 CC
 SQ Sequence 327 AA;
 Query Match 44.6%; Score 749; DB 2; Length 327;
 Best Local Similarity 47.3%; Pred. No. 1.2e-53;
 Matches 156; Conservative 63; Mismatches 105; Indels 6; Gaps 2;
 QY 8 KLANVLLATGTTAGAGASANSATYQAAKVGVDKLIAGVPELADLANVREGVMOIAS 67
 DB 3 KLPNVLLATGTTAGASATGTQTTGYAGALGVDTLLNAVPEVKLANVGEOPSNMAS 62
 QY 68 ESITNDLLKLASSVAELADSNVDGIVITHGTDLSEETAFPLNVEKTPVIVVGS 127
 DB 63 ENMTGDVLLKSKRVNELAADVDGIVITHGTDLSEETAFPLNVEKTPVIVVGS 122
 QY 128 PGTMSADGMLNVAVAVASNKDSRGKGLVTMNDIEQSDVSKSINIKTEAFK-SAW 186
 DB 123 PATASADGPNMLLEAVAVAGDQSRGKGVVTVNDRIASARYITKTNASTLDFEKABE 182
 QY 187 GPLGVVSGSYWFRLLPAKHTVNSSEPDIKOISLPQVDIAYSGNVTDTAYKALQNGA 246
 DB 193 GYLGVITGNRIYQNRIDKHTTTSVFPVRGLTSLPKVDILYGYQDDPEVLYDAIYQGV 242
 QY 247 KALIHAGTNGSVSSRVVPLQELRKNGVQIIRSSRQGGFVLRNAEQPDDKNDVVYAH 306
 DB 243 KGIYVAGNGAGSVSRGIRGAKRKAMEKGVVIRSTRNGIV-----PPDELPGLVSDS 297
 QY 307 LNPOKARILLAVAMTKTQDSKELORIEMV 336
 DB 298 LNPAPARILLALTRTSDPKVIOEYFTY 327
 RESULT 7
 AAP70348
 ID AAP70348 standard; protein, 348 AA.
 AC AAP70348;
 XX
 XX 25-MAR-2003 (revised)
 DT 07-AUG-1991 (first entry)
 XX
 XX Sequence of a protein with L-asparaginase (LA) activity in PASN 30 or
 DE PASN 32.
 DE
 XX Enzyme; cancer therapy; leukaemia therapy.
 KW

```

XX OS Erwinia chrysanthemi.
XX FH Key Location/Qualifiers
XX FT Peptide 1..21
XX FT /Label= signal
XX FT Protein 22..348
XX PN EP211639-A.
XX PD 25-FEB-1987.
XX PF 04-AUG-1986; 86EP-00305984.
XX PR 06-AUG-1985; 85GB-00019753.
XX PA (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
XX PI Atkinson A, Manton NP, Gilbert HJ;
XX PI MPI, 1987-051784/08.
XX DR N-PSDB; AAN70557.
XX PT New recombinant plasmids coding for L-asparaginase - esp. from Erwinia
XX PT carotovora, useful for treating malignancies and new transformed hosts.
XX PS Disclosure, Fig 4, 35pp; English.
XX CC When E.carotovora SCRI 193 contg. plasmid pAGN 32 was cultured, the cell-
XX CC free culture broth contained 11.19 LA units/ml with specific activity
XX CC 53.34 units/mg of protein. These figures compare with 10.87 LA units/ml
XX CC and 14.97 units/mg of protein for the currently used prodn. strain
XX CC E.chrysanthemi NCPB 1066. (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 348 AA;

Query Match 44.6%; Score 749; DB 1; Length 348;
Best Local Similarity 47.3%; Pred. No. 1.3e-53;
Matches 156; Conservative 63; Mismatches 105; Indels 6; Gaps 2;

QY 8 KLANVITLTTGGTIGAGSAAANSATYQAKVGVNKLINGVPELADIANVREGQVQIAS 67
DB 24 KLPNIVITLTTGGTIGAGSAAATGTGTGKAGLVDTLINAVPEVKLANVKGQFENMAS 83
QY 68 ESTINDDLKLAASVAELADSDVDGIVTTHGDTLEETAYFLNLTVEKTDKPIVVGSMR 127
DB 84 ENMTGIVVLKLSQRYNELRLARDVDGIVTTHGDTVEESAYFLHLTVKSDKPVVFAAMR 143
QY 128 PGTASADGMLNLNNAVAASNKDSRGKGLVYTMNDEIOSGRDVSINIKTEAFK-SAM 186
DB 144 PATASADGPMNLLEAVRVAAGDKQSRGVMVVLNDRIGSARTITKTNASTLDTFFANEE 203
QY 187 GPLGMVBEKSYWFLPAKHTVNSEFDIKQISLFPQVDIAYSYGAVTDTAYKALAQNGA 246
DB 204 GYLGVITIGRIYQNNIDKLTHTRSVFDVRGLTSLPKVOLLVGYQDDPELVDAALQHGCV 263
QY 247 KALIHAGTNGSVSSRVVPALQELRNQVQIIRSSQGGGFVLRNABQDDKNDVVAHD 306
DB 264 KGIIVYAGMGAGSVSVGIAIGMRKAMEKGVVIRSTRGTNGIV-----PDEBELPGLVSDS 318
QY 307 LNPQKARILLAMVAMTKQDSKEIQRIFWEY 336
DB 319 LNPQKARILLMLALTRTSDPKVIQGEYFHTY 348

RESULT 8
ABP98617 standard; protein; 348 AA.
AC ABP98617,
XX
XX 30-MAY-2003 (first entry)
XX

```

```

DB OS erwinia chrysanthemi L-asparaginase protein.
XX KM Asparaginase; cytosolic; blood; acute lymphoblastic leukemia.
XX OS Erwinia chrysanthemi.
XX OS WO2003018742-A2.
XX PN 06-MAR-2003.
XX PD 21-AUG-2002; 2002MO-RU000405.
XX PF 22-AUG-2001; 2001RU-00123442.
XX PR 04-APR-2002; 2002RU-00108505.
XX PA (UCHR=) GOS UCHREZHDENIE NI INST BIOMED KHMITI.
XX PA (UCHR=) GOS UCHREZHDENIE TSENTR BIOINZHEMERIYA.
XX PI Eldarov MA, Zhgun AA, Gervaziev YV, Aleksandrova SS, Bogush VG;
XX PI Sidoruk KV, Sveshnikova EV, Borisova AA, Omeinyuk NM, Archakov AI;
XX PI Skryabin KG, Sokolov NN;
XX DR MPI; 2003-268418/26.
XX PT New isolated Erwinia carotovora L-asparaginase gene useful for producing
XX PT a L-asparaginase protein useful for treating malignant and nonmalignant
XX PT diseases of the blood.
XX PS Example 2; Page 9-10; 29pp; Russian.
XX XS This sequence represents the complete form of an Erwinia chrysanthemi L-
XX CC asparaginase. The invention relates to the isolation of the gene and
XX CC protein for L-asparaginase from Erwinia carotovora and to the production
XX CC of the mature version of the protein especially by expression of the DNA
XX CC in a bacterial strain and has low affinity for L-glutamine. The protein
XX CC is useful for treating malignant and non-malignant diseases of the blood
XX CC (e.g. acute lymphoblastic leukemia)
XX SQ Sequence 348 AA;

Query Match 44.6%; Score 749; DB 6; Length 348;
Best Local Similarity 47.3%; Pred. No. 1.3e-53;
Matches 156; Conservative 63; Mismatches 105; Indels 6; Gaps 2;

QY 8 KLANVITLTTGGTIGAGSAAANSATYQAKVGVNKLINGVPELADIANVREGQVQIAS 67
DB 24 KLPNIVITLTTGGTIGAGSAAATGTGTGKAGLVDTLINAVPEVKLANVKGQFENMAS 83
QY 68 ESTINDDLKLAASVAELADSDVDGIVTTHGDTLEETAYFLNLTVEKTDKPIVVGSMR 127
DB 84 ENMTGIVVLKLSQRYNELRLARDVDGIVTTHGDTVEESAYFLHLTVKSDKPVVFAAMR 143
QY 128 PGTASADGMLNLNNAVAASNKDSRGKGLVYTMNDEIOSGRDVSINIKTEAFK-SAM 186
DB 144 PATASADGPMNLLEAVRVAAGDKQSRGVMVVLNDRIGSARTITKTNASTLDTFFANEE 203
QY 187 GPLGMVBEKSYWFLPAKHTVNSEFDIKQISLFPQVDIAYSYGAVTDTAYKALAQNGA 246
DB 204 GYLGVITIGRIYQNNIDKLTHTRSVFDVRGLTSLPKVOLLVGYQDDPELVDAALQHGCV 263
QY 247 KALIHAGTNGSVSSRVVPALQELRNQVQIIRSSQGGGFVLRNABQDDKNDVVAHD 306
DB 264 KGIIVYAGMGAGSVSVGIAIGMRKAMEKGVVIRSTRGTNGIV-----PDEBELPGLVSDS 318
QY 307 LNPQKARILLAMVAMTKQDSKEIQRIFWEY 336
DB 319 LNPQKARILLMLALTRTSDPKVIQGEYFHTY 348

RESULT 9
AAW98609 standard; protein; 364 AA.
XX
XX
XX

```

AC AAW98609;
 XX
 DT 31-MAR-1999 (first entry)
 XX
 DE H. pylori GHPO 187 protein.
 XX
 KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
 XX peptic ulcer disease.
 XX
 OS Helicobacter pylori.
 XX
 FN W09843478-A1.
 XX
 PD 08-OCT-1998.
 XX
 PF 01-APR-1998; 98WO-US006371.
 XX
 PR 01-APR-1997; 97US-00833457,
 XX 24-JUN-1997; 97US-00881227,
 PR 29-JUL-1997; 97US-00902615.
 XX
 PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 XX (HUMAN-) HUMAN GENOME SCI INC.
 XX
 PI Kleantious H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
 XX
 DR MPI; 1998-542293/46.
 XX N-PSDB; AAX14328.
 XX
 PT New isolated Helicobacter polymucleotides - used to develop products for
 XX the diagnosis, prevention and treatment of Helicobacter infections and
 PT gastrointestinal diseases.
 XX
 PS Claim 8; Page 1227-1229; 2054pp; English.
 XX
 CC This sequence represents a Helicobacter pylori GHPO protein of the
 CC invention. The polypeptides can be used for preventing or treating
 CC Helicobacter infections, and gastroduodenal diseases associated with
 CC these infections, including acute, chronic, and atrophic gastritis, and
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
 CC used for the production of antibodies. The products can also be used for
 CC detection and diagnosis
 CC
 SQ Sequence 364 AA;
 Query Match 42.7%; Score 717; DB 2; Length 364;
 Best Local Similarity 44.6%; Pred. No. 6.1e-51;
 Matches 148; Conservative 75; Mismatches 103; Indels 6; Gaps 5;
 QY 7 QKLANVVIATGTTAGAGASANSATYQAAKVGVDKLAGVPELADLANVGEQVMQIA 66
 DB 37 QNLPTALATGTTAGTSGANS-ASLSGYSKSGELGKELKAIPLNRLARIGEQISNIG 95
 QY 67 SSSITNDLLKLAASSVAELADSNVDGIVYTHGTDTLEETAYFLNLVEKTRPIYVGS 126
 DB 96 SDDMEERWFKAKRAQELLDSDRIGQVYTHGTDTLEESAYFLNLVLRSTRPVVLVGM 155
 QY 127 RRGTMASADGMLNVAVVAASNDKSRGKGLVTMDEIOSGRDVSXINITEAFKSA- 185
 DB 156 RRAASISADGALNVAASVVALNEKSAKNGVAVLMDNIFSAEYVKTHTTSTFKALN 215
 QY 186 WGPLGAVVEGKSYWFLPAKRHTVNSEPDIKOISS-LPOVDIAYSXGVTDTAYKALAQ 244
 DB 216 SGATSGSVYVGTQRYVWQPLRKHTTESSEFSLQKTPLPKVDIITHTAGMTPLPQASLNS 275
 QY 245 GAKALIHAGTNGSVSSRVVPALQELRKNGVQITRSSRQGGFTLRNAEQPDDKNDWYA 304
 DB 276 HAKGVVIAVGNGSVSAGFLKAMQEASQMGVIVRSRVSGBI--TSGEIDK-AFITS 332
 QY 305 HDLNPOKARILLAMVAMTKTQDSKELQRIFWEX 336
 DB 333 DNLNPOKARILLQALTLTKTNKKEKIQEMPEEY 364

RESULT 10
 ABO61384
 ID ABO61384 standard; protein; 356 AA.
 XX
 AC ABO61384;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Klebsiella pneumoniae polypeptide seqid 7901.
 XX
 KW Recombinant expression vector; transcription regulatory element;
 XX Klebsiella pneumoniae protein; antibacterial; vaccine.
 XX
 OS Klebsiella pneumoniae.
 XX
 FN US6610836-B1.
 XX
 PD 26-AUG-2003.
 XX
 PF 27-JAN-2000; 2000US-00489039.
 XX
 PR 29-JAN-1999; 99US-0117747P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton GL, Osborne M;
 XX
 DR MPI; 2003-895346/82.
 XX N-PSDB; ACH94935.
 XX
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 XX preparing a vaccine composition against Klebsiella pneumoniae.
 PT
 XX
 PS Disclosure; SEQ ID NO 7901; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 CC
 SQ Sequence 356 AA;
 Query Match 42.3%; Score 711; DB 7; Length 356;
 Best Local Similarity 45.0%; Pred. No. 1.9e-50;
 Matches 150; Conservative 68; Mismatches 109; Indels 6; Gaps 2;
 QY 5 NQOKLANVVIATGTTAGAGASANSATYQAAKVGVDKLAGVPELADLANVGEQVMQ 64
 DB 29 SETRLPHVILLATGTTAGTSGAASNTQGTGYKAGAGVOTLINAVPEMSKIAHVEGEQVAN 88
 QY 65 IASISITNDLLKLAASSVAELADSNVDGIVYTHGTDTLEETAYFLNLVEKTRPIYV 124
 DB 89 IGSEMTSDIILQSKRVNALLARDVDGIVYTHGTDTLEETAYFLNLVTSNKEVFTFA 148
 QY 125 SMRPGTMASADGMLNVAVVAASNDKSRGKGLVTMDEIOSGRDVSXINITEAFKSA 184
 DB 149 AARFPTALISADGPMMLLEAVTVVADPDARGSGVWVLDNRIGAAAFVTKTNATSLDTERA 208
 QY 185 -AMGPLGAVVEGKSYWFLPAKRHTVNSEPDIKOISS-LPOVDIAYSXGVTDTAYKALAQ 243
 DB 209 PEEGYLVGVGKPGQFERVRVDKIHRLRSEFVDRQKVLPKVYIIGYQDDPEYMDAALA 268
 QY 244 NGAKALIHAGTNGSVSSRVVPALQELRKNGVQITRSSRQGGFTLRNAEQPDDKNDWY 303
 DB 269 HHADGIYAGTGAGSVRSAAIGIKAAQOAGIVVVRASRTSGGV-----PPDDSQPGIV 323
 QY 304 AHDLPKARILLAMVAMTKTQDSKELQRIFWEX 336
 DB 324 ADSINPAPARILLMTALVOTKDPOLIOQYFHTY 356

RESULT 11
ID ADP05106 standard; protein, 346 AA.
XX
AC ADP05106;
XX
XX 12-FEB-2004 (first entry)
DT
DE Bacterial polypeptide #1219.
XX
XX Proteus mirabilis infection; bacterial infection; antibacterial;
KW immunostimulant.
XX
OS Proteus mirabilis.
XX
PN US6605709-B1.
XX
PD 12-AUG-2003.
XX
PF 05-APR-2000; 2000US-00543681.
XX
PR 09-APR-1999; 99US-0128706P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
XX WPI; 2003-895291/82.
DR N-PSDB; ADF00934.
XX
XX New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX
XX Disclosure; SEQ ID NO 5391; 870pp; English.
XX
XX The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polypeptide of the invention.
XX
SQ Sequence 346 AA;
Query Match 42.3%; Score 710.5; DB 7; Length 346;
Best Local Similarity 46.2%; Pred. No. 2e-50;
Matches 152; Conservative 58; Mismatches 114; Indels 5; Gaps 4;

9 LANVVLATGCTIAGAGASANSATYQAAKVGVDKLIAGVPELADLANVGEQWQIASE 68
22 LBNVTLATGCTIAGAGGDS-ATSSSYTAGLGIIDLINNVEAKKYANLTGSEVNVISQ 80
69 SITNDLKLASSVAELADSNVDGIVITHGDTLTETNAYFLNLTVEKTKPIYVVGSKRP 128
81 DNMDDWMLKANKIN--TTCDDKTDGVIITHTGTIMERTAYFLDLTTACKKPPVVMVGAMP 138
129 GTAMSDGMLNTYNAAVANSNDKSRGKGLVTNNDLIGSRDYSKSIKTBKFSKA-WG 187
139 ATRLGADGFLANLNAVAVASDKASENRGVALVTNNSVISGKDDVKNNTTEVQAFQINAG 198
188 PLGMVVEGSKYWFRLPAKHTVNSEPDIIQISLPOVDIAYSIGNVTDPAKALQNGAK 247
199 AGGYVADGKAVHYTYATTPR-ABKVAFDVGSLTELPLFVGLVYNNVANSDDLPAKAFINHRK 257

246 ALIHAGTNGSVSRVVPALQELRKNGVOIIRSRQGGFVLNNAEQPDQKDWVVAHDL 307
258 GIVSAGVNGNLYSDILNTLADGKGVVAVSSSRVPVFTTONGEVDARKGFIASERL 317
308 NPQAKRIILAMVAMTKTQDSKELORIEMEX 336
318 NPQAKRVLILQLSTETQDPATIQENPEKY 346
Db

RESULT 12
ABM68938
ID ABM68938 standard; protein, 348 AA.
XX
AC ABM68938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Photorhabdus luminescens protein sequence #2035.
XX
XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough.
XX
OS Photorhabdus luminescens.
XX
PN WO200294867-A2.
XX
XX 28-NOV-2002.
XX
PF 07-FEB-2002; 2002WO-IB003040.
XX
PR 07-FEB-2001; 2001FR-00001659.
XX
PA (INSP) INST PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Duchaud E, Taucrit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI Buchrieser C;
XX
XX WPI; 2003-148459/14.
XX
PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX
XX Claim 2; SEQ ID NO 2035; 1205pp; French.
XX
XX The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens proteins
XX
SQ Sequence 348 AA;
Query Match 40.8%; Score 685.5; DB 6; Length 348;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 21:25:54 / Search time 43 seconds
(without alignments)
583.304 Million cell updates/sec

Title: US-09-842-628-2

Perfect score: 1680

Sequence: 1 KEVENQOKLVAVILATGST.....MVAMTKTQDSKELQRIFWRY 336

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCUTS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	1680	100.0	336	5	PCT-US92-10421-2
2	1646.5	98.0	337	3	US-08-050-482A-2
3	1432.5	85.3	333	4	US-09-252-991A-22364
4	1063.5	63.3	374	4	US-09-328-352-6462
5	749	44.6	327	1	US-07-867-105B-1
6	711	42.3	356	4	US-09-489-039A-7901
7	710.5	42.3	336	4	US-09-543-681A-5391
8	372.5	22.2	333	4	US-09-134-000C-4613
9	361.5	21.5	322	4	US-09-710-279-1912
10	361.5	21.5	326	3	US-09-134-001C-3234
11	356	21.2	320	4	US-09-583-110-3586
12	356	21.2	321	4	US-09-107-433-3394
13	331.5	19.7	334	4	US-09-107-532A-4947
14	194	11.5	347	4	US-09-543-681A-8255
15	191.5	11.4	227	4	US-09-248-796A-17491
16	183	10.9	385	4	US-09-489-039A-7667
17	160.5	9.6	530	2	US-08-329-501-12
18	160.5	9.6	530	3	US-09-140-177-12
19	160.5	9.6	530	3	US-09-397-979-12
20	154	9.2	506	2	US-08-329-501-12
21	154	9.2	506	3	US-09-140-177-2
22	154	9.2	506	3	US-09-397-979-2
23	153.5	9.1	354	1	US-08-222-289-2
24	152	9.0	332	4	US-09-252-991A-25262
25	145.5	8.7	363	3	US-09-195-666A-4
26	145.5	8.7	363	3	US-09-635-705-4
27	145.5	8.7	363	4	US-09-634-858A-4

28	145.5	8.7	363	4	US-08-869-927C-4	Sequence 4, Appl1
29	145.5	8.7	565	3	US-09-008-481A-6	Sequence 6, Appl1
30	145.5	8.7	565	3	US-09-195-666A-5	Sequence 5, Appl1
31	145.5	8.7	565	3	US-09-195-666A-49	Sequence 49, Appl1
32	145.5	8.7	565	3	US-09-309-592-6	Sequence 6, Appl1
33	145.5	8.7	565	3	US-09-635-705-5	Sequence 5, Appl1
34	145.5	8.7	565	3	US-09-635-705-49	Sequence 49, Appl1
35	145.5	8.7	565	4	US-09-634-858A-5	Sequence 5, Appl1
36	145.5	8.7	565	4	US-08-869-927C-49	Sequence 49, Appl1
37	145.5	8.7	565	4	US-08-869-927C-5	Sequence 5, Appl1
38	145.5	8.7	565	4	US-08-869-927C-49	Sequence 49, Appl1
39	143	8.5	339	4	US-09-902-540-15220	Sequence 15220, A
40	126	7.5	365	3	US-09-195-666A-6	Sequence 6, Appl1
41	126	7.5	365	3	US-09-195-666A-7	Sequence 7, Appl1
42	126	7.5	365	3	US-09-635-705-6	Sequence 6, Appl1
43	126	7.5	365	3	US-09-635-705-7	Sequence 7, Appl1
44	126	7.5	365	4	US-09-634-858A-6	Sequence 6, Appl1
45	126	7.5	365	4	US-09-634-858A-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
PCT-US92-10421-2

Sequence 2, Application PC/TUS9210421

GENERAL INFORMATION:

APPLICANT: Robert, Joseph

APPLICANT: MacCallister, Thomas W

APPLICANT: Sethuraman, Natrajan

APPLICANT: Freeman, Abbie G

TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND

TITLE OF INVENTION: ITS USE IN ANTIVIRAL AND ANTICANCER THERAPY

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESSES:

ADDRESSER: Banner, Birch, McKie and Beckett

STREET: 1001 G Street N.W.

CITY: Washington, D.C.

COUNTRY: U.S.A.

ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/10421

FILING DATE: 19921204

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P4140003.8

FILING DATE: 04-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A

REGISTRATION NUMBER: 32,141

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 336 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULAR TYPE: protein

HYPOTHETICAL: YES

ANTI-SENSE: NO

PCT-US92-10421-2

Query Match 100.0%; Score 1680; DB 5; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.9e-149; Indels 0; Gaps 0;
Matches 336; Conservative 0; Mismatches 0;

QY 1 KEVENOQKLANVVIATGTTAGAGASANSATYQAAYGVDKLIAGVPELADLANVGE 60
DB 1 KEVENOQKLANVVIATGTTAGAGASANSATYQAAYGVDKLIAGVPELADLANVGE 60
QY 61 QVMOJASESTINDDLKLKASSVAELADSDVDGIYTHGTDTLEETAFVFLNVEKTDKPI 120
DB 61 QVMOJASESTINDDLKLKASSVAELADSDVDGIYTHGTDTLEETAFVFLNVEKTDKPI 120
QY 121 VVGSMPGTAMSDGMLNLVNAVAVASNKDSRGKGVLTWNDEIQSGRDVSKSINIKTE 180
DB 121 VVGSMPGTAMSDGMLNLVNAVAVASNKDSRGKGVLTWNDEIQSGRDVSKSINIKTE 180
QY 181 AFKSAWGLGAVVEGKSYWFLPAKRHTVNSEFDIKQISSLPQVDIAYSNGVTTAYKA 240
DB 181 AFKSAWGLGAVVEGKSYWFLPAKRHTVNSEFDIKQISSLPQVDIAYSNGVTTAYKA 240
QY 241 LAONGAKALIHAGTNGSVSSRVVPALQELRNKGVOIIRSSHVNGGQGVFLRNAEQPDDKN 300
DB 241 LAONGAKALIHAGTNGSVSSRVVPALQELRNKGVOIIRSSHVNGGQGVFLRNAEQPDDKN 300
QY 301 WVAHADLNPOKARILANVAMTKTODSKELQRIFWEX 336
DB 301 WVAHADLNPOKARILANVAMTKTODSKELQRIFWEX 336

RESULT 2

US-08-050-482A-2
Sequence 2, Application US/08050482A
Patent No. 6512939

GENERAL INFORMATION:
APPLICANT: ROBERTS, Joseph
MACALISTER, Thomas W.
SETHURAMAN, Natarajan

TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND
FREEMAN, Abbie G.

NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSER: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.

COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/050,482A
FILING DATE: 25-Apr-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US92/10421
FILING DATE: 04-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 023032/0106
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 337 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-08-050-482A-2

Query Match 98.0%, Score 1646.5, DB 3, Length 337,

Best Local Similarity 98.2%, Pred. No. 2.6e-146;
Matches 331; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1 KEVENOQKLANVVIATGTTAGAGASANSATYQAAYGVDKLIAGVPELADLANVGE 60
DB 1 KEVENOQKLANVVIATGTTAGAGASANSATYQAAYGVDKLIAGVPELADLANVGE 60
QY 61 QVMOJASESTINDDLKLKASSVAELADSDVDGIYTHGTDTLEETAFVFLNVEKTDKPI 120
DB 61 QVMOJASESTINDDLKLKASSVAELADSDVDGIYTHGTDTLEETAFVFLNVEKTDKPI 120
QY 121 VVGSMPGTAMSDGMLNLVNAVAVASNKDSRGKGVLTWNDEIQSGRDVSKSINIKTE 180
DB 121 VVGSMPGTAMSDGMLNLVNAVAVASNKDSRGKGVLTWNDEIQSGRDVSKSINIKTE 180
QY 181 AFKSAWGLGAVVEGKSYWFLPAKRHTVNSEFDIKQISSLPQVDIAYSNGVTTAYKA 240
DB 181 AFKSAWGLGAVVEGKSYWFLPAKRHTVNSEFDIKQISSLPQVDIAYSNGVTTAYKA 240
QY 241 LAONGAKALIHAGTNGSVSSRVVPALQELRNKGVOIIRSSHVNGGQGVFLRNAEQPDDKN 299
DB 241 LAONGAKALIHAGTNGSVSSRVVPALQELRNKGVOIIRSSHVNGGQGVFLRNAEQPDDKN 300
QY 300 DWVAHADLNPOKARILANVAMTKTODSKELQRIFWEX 336
DB 301 DWVAHADLNPOKARILANVAMTKTODSKELQRIFWEX 337

RESULT 3

US-09-252-991A-22364
Sequence 22364, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22364
LENGTH: 393
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22364

Query Match 85.3%, Score 1432.5, DB 4, Length 393;
Best Local Similarity 84.3%, Pred. No. 4.3e-126;
Matches 284; Conservative 22; Mismatches 30; Indels 1; Gaps 1;

QY 1 KEVENOQKLANVVIATGTTAGAGASANSATYQAAYGVDKLIAGVPELADLANVGE 60
DB 57 KEVENOQKLANVVIATGTTAGAGASANSATYQAAYGVDPQLASVPQDKIDIANVGE 116
QY 61 QVMOJASESTINDDLKLKASSVAELADSDVDGIYTHGTDTLEETAFVFLNVEKTDKPI 120
DB 117 QVMOJASESTINDDLKLKASSVAELADSDVDGIYTHGTDTLEETAFVFLNVEKTDKPI 176
QY 121 VVGSMPGTAMSDGMLNLVNAVAVASNKDSRGKGVLTWNDEIQSGRDVSKSINIKTE 180
DB 121 VVGSMPGTAMSDGMLNLVNAVAVASNKDSRGKGVLTWNDEIQSGRDVSKSINIKTE 180
QY 181 AFKSAWGLGAVVEGKSYWFLPAKRHTVNSEFDIKQISSLPQVDIAYSNGVTTAYKA 240
DB 237 AFKSAWGLGAVVEGKSYWFLPAKRHTVNSEFDIKQISSLPQVDIAYSNGVTTAYKA 296
QY 241 LAONGAKALIHAGTNGSVSSRVVPALQELRNKGVOIIRSSHVNGGQGVFLRNAEQPDDKN 299
DB 297 LAONGAKALIHAGTNGSVSSRVVPALQELRNKGVOIIRSSHVNGGQGVFLRNAEQPDDKN 356

QY 300 DNVVAHDLNPKARILAMVAMTKTODSKELORIFMEY 336
 DB 357 DNVVAHDLNPKARILAMVAMTKTODSKELORIFMEY 393

RESULT 4
 US-09-328-352-6462
 Sequence 6462, Application US/09328352
 Patent No. 6562958
 GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC99-03PA
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 6462
 LENGTH: 374
 TYPE: PRT
 ORGANISM: Acinetobacter baumannii
 US-09-328-352-6462

Query Match 63.3%; Score 1063.5; DB 4; Length 374;
 Best Local Similarity 64.7%; Pred. No. 1,9e-91;
 Matches 213; Conservative 45; Mismatches 68; Indels 3; Gaps 2;

QY 11 NVVILATGTTAGAGSAANSATYQAAKVGVDKLINGVELDLANRGEQVMQIASSEI 70
 DB 46 NVVVATGTTAGAGSANSATYQAAKVPDALINAVQIOLIANVSGIQALQVASSSEI 105
 QY 71 TMDDLKLASSVAELADSDNDGIVTHTGDTLEETAYFLINVEKTDKRIYVVGSRPST 130
 DB 106 TDKELQARQVNEIVKPKTVNGVITHTGDTLEETAYFLINVEKTDKRIYVVGSRPST 165
 QY 131 AMSADGMLNLYNAVAVASNKDSRGKGLVTNDEIOSGSDVSKSINIKTEAFKSAWGPJG 190
 DB 166 ALSADGPMNLVSAVALAASDDAKNKGVWTLNDSIFAADVTGKGINIHNAVVSQMGALG 225
 QY 226 TLVEGKPYFRQSVKHTVASSENIENIKGDALPTVQIYVGSQSMLEPAYEAVAKAGAYA 285
 DB 249 LIHAGTNGSVSSRVVPALQELR-KKGVQIIRSSRQGGFYLRNAEQPDKDKMDVVAHDL 307
 QY 286 IHAAGTNGSVAKYVPTLQNLHDKNGIQIIRSSRVPGFVLDAEQPDSPKYGWAHDL 345
 QY 308 NPKARILAMVAMTKTODSKELORIFMEY 336
 DB 346 NPKARILAMVAMTKTODSKELORIFMEY 374

RESULT 5
 US-07-867-105B-1
 Sequence 1, Application US/07867105B
 Patent No. 5310670
 GENERAL INFORMATION:
 APPLICANT: GOMARD, Christopher R.
 TITLE OF INVENTION: Improved Method for the
 TITLE OF INVENTION: Purification of Erwinia
 TITLE OF INVENTION: L-asparaginase
 NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Kirscheitein, Octinger, Israel
 ADDRESSES: & Schifftmiller, P.C.
 STREET: 551 Fifth Avenue
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: NY 10176-0024
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 5.25 inch diskette

COMPUTER: Olivetti (IBM Compatible)
 OPERATING SYSTEM: MS DOS
 SOFTWARE: NLI
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/867,105B
 FILING DATE: 19920625
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9017002.8
 FILING DATE: 02-AUG-90
 ATTORNEY/AGENT INFORMATION:
 NAME: David B. Kirscheitein
 REGISTRATION NUMBER: 17244
 REFERENCE/DOCKET NUMBER: -
 NAME: Alan Israel
 REGISTRATION NUMBER: 27564
 REFERENCE/DOCKET NUMBER: -
 NAME: Martin W. Schifftmiller
 REGISTRATION NUMBER: 30421
 REFERENCE/DOCKET NUMBER: -
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 327 a.a.
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: Linear
 MOLECULE TYPE: Polypeptide
 DESCRIPTION: The amino acid sequence of Erwinia
 DESCRIPTION: chryseanthemi L-asparaginase
 HYPOTHEICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Erwinia chryseanthemi
 US-07-867-105B-1

Query Match 44.6%; Score 749; DB 1; Length 327;
 Best Local Similarity 47.3%; Pred. No. 5.4e-62;
 Matches 156; Conservative 63; Mismatches 105; Indels 6; Gaps 2;

QY 8 KLANVILATGTTAGAGSAANSATYQAAKVGVDKLINGVELDLANRGEQVMQIASSEI 67
 DB 3 KLPNVILATGTTAGAGSAANSATYQAAKVGVDKLINGVELDLANRGEQVMQIASSEI 62
 QY 68 ESITMDDLKLASSVAELADSDNDGIVTHTGDTLEETAYFLINVEKTDKRIYVVGSRPST 127
 DB 63 EMTGDTVLKLSQRNHELARDVDGIVTHTGDTLEETAYFLINVEKTDKRIYVVGSRPST 122
 QY 128 PGTASADGMLNLYNAVAVASNKDSRGKGLVTNDEIOSGSDVSKSINIKTEAFKSAWGPJG 186
 DB 123 PATASADGPMNLVSAVALAASDDAKNKGVWTLNDSIFAADVTGKGINIHNAVVSQMGALG 182
 QY 187 GPLGAVVEKSYWRLPAKHTVNSEFDIKQISLPOVDIAYSXGVTDTAYKALAONGA 246
 DB 183 GYLGVIGIRIYQNRIDGLATRTSVFDRGLTSLPKVDILYGYDDPELYLDAALQGV 242
 QY 247 KALIHAGTNGSVSSRVVPALQELRKNQVQIIRSSRQGGFYLRNAEQPDKDKMDVVAHDL 306
 DB 243 KGIIVAGKAGSVSRVGLAGKAKMEKGVVIRSTRGTNGIY-----PDDELPGLVSDS 297
 QY 307 INPKARILAMVAMTKTODSKELORIFMEY 336
 DB 298 INPKARILAMVAMTKTODSKELORIFMEY 327

RESULT 6
 US-09-489-039A-7901
 Sequence 7901, Application US/09489039A
 Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gary Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709.2004001
 CURRENT APPLICATION NUMBER: US/09/489,039A

/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 7901
/ LENGTH: 356
/ TYPE: PRT
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7901

Query Match 42.3%; Score 711; DB 4; Length 356;
Best Local Similarity 45.0%; Pred. No. 2.3e-58;
Matches 150; Conservative 68; Mismatches 109; Indels 6; Gaps 2;

QY 5 NQOKLAVVILATGGTIGAGASAAANSATYQAKVGVDTLIGVPELADLANVREGVQWQ 64
DB 29 SETRLPHVILATGGTIGAGASAAANSATYQAKVGVDTLIGVPELADLANVREGVQWQ 88
QY 65 IASEITNDLKLASSVAELADSDVDGIVITHTGDTLEETAYFLNVEKTDKPIVVG 124
DB 89 IGSEMTSTIITLQSKRVALLARDVDGIVITHTGDTLEETAYFLNVEKTDKPIVVG 148
QY 125 SMREGTASADGMLNNAVAVASNDKSGKVLVTMDEIOSGRDVSSTIKTEAPRS 184
DB 149 AMRPATASADGPMNLLEAVTVAADPDARGGVAVVLDRIGAARFVTKNTSLDTPRA 208
QY 185 -AMGPIAGWVEGSKYWFRLPAKRHTVNSEFDIKQISLPQVDIAYSQVNTDTAYKALAQ 243
DB 209 PEEGLIGVVGKPGQEPETRVVDKHTLRSVFVDRQLKVLPRVITIGYDDPEYMTDAIA 268
QY 244 NGAKALIHAGTNGSVSRVFPALQELRNQVOIIRSRHQGGFYLRNAEOPDDKNDWV 303
DB 269 HHADGIVYAGTSGSVSAAGIKKQAGIIVVRASTRSGV-----PPDDSGPGLV 323
QY 304 AMDLNPORARILAMVAMTQDSKELQRIEMV 336
DB 324 ADLNPORARILAMVAMTQDSKELQRIEMV 356

RESULT 7

US-09-543-681A-5391
/ Sequence 5391, Application US/09543681A
/ Patent No. 6605709
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ FILE REFERENCE: 2709.1002-001
/ CURRENT APPLICATION NUMBER: US/09/543,681A
/ CURRENT FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 60/128,706
/ PRIOR FILING DATE: 1999-04-09
/ NUMBER OF SEQ ID NOS: 8344
/ SEQ ID NO 5391
/ LENGTH: 346
/ TYPE: PRT
/ ORGANISM: Proteus mirabilis
US-09-543-681A-5391

Query Match 42.3%; Score 710.5; DB 4; Length 346;
Best Local Similarity 46.2%; Pred. No. 2.4e-58;
Matches 152; Conservative 58; Mismatches 114; Indels 5; Gaps 4;

QY 9 LANVYILATGGTIGAGASAAANSATYQAKVGVDTLIGVPELADLANVREGVQWQ 68
DB 22 LPNVYILATGGTIGAGASAAANSATYQAKVGVDTLIGVPELADLANVREGVQWQ 80
QY 69 SITNDLKLASSVAELADSDVDGIVITHTGDTLEETAYFLNVEKTDKPIVVGSMRP 128
DB 81 DMNDQVWLKLANKIN--TDCDKTIDGVIHTGDTLEETAYFLNVEKTDKPIVVGSMRP 138
QY 129 GTASADGMLNNAVAVASNDKSGKVLVTMDEIOSGRDVSSTIKTEAPRS-WG 187

DB 139 ATALGADPELNLNNAVAVASNDKSGKVLVTMDEIOSGRDVSSTIKTEAPRS 198
QY 188 PLGAVVEGSKYWFRLPAKRHTVNSEFDIKQISLPQVDIAYSQVNTDTAYKALQNGAK 247
DB 199 AOGYVHDKVHTYTRATR-AEKVAFVSKTELPKQIVNRYANASDLPAKAFIDNFK 257
QY 248 ALIHAGTNGSVSRVFPALQELRNQVOIIRSRHQGGFYLRNAEOPDDKNDWV 307
DB 258 GIVSAGVNGNLVSDILNTLADGVKGVVVRSSRPVFTQNGRVDADKGFASBRL 317
QY 308 NPORARILAMVAMTQDSKELQRIEMV 336
DB 318 NPORARILAMVAMTQDSKELQRIEMV 346

RESULT 8

US-09-134-000C-4613
/ Sequence 4613, Application US/09134000C
/ Patent No. 6617156
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
/ FILE REFERENCE: 032796-032
/ CURRENT APPLICATION NUMBER: US/09/134,000C
/ CURRENT FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/055,778
/ PRIOR FILING DATE: 1997-08-15
/ NUMBER OF SEQ ID NOS: 6812
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 4613
/ LENGTH: 333
/ TYPE: PRT
/ ORGANISM: Enterococcus faecalis
US-09-134-000C-4613

Query Match 22.2%; Score 372.5; DB 4; Length 333;
Best Local Similarity 30.0%; Pred. No. 1.3e-26;
Matches 105; Conservative 67; Mismatches 137; Indels 41; Gaps 11;

QY 4 ENQOKLAVVILATGGTIA---GAGASAA---NSATYQAKVGVDTLIGVPELADLAN 56
DB 3 QRRKOMKTLVLTHTGTTISMSKSVAGVAPEKPLMEQEA-----LPSGVHLV--- 52
QY 57 VRGEVMOIASEITNDLKLASSVAELADSDVDGIVITHTGDTLEETAYFLNVEKTD 116
DB 53 --VEDIFRIPPHMTLEBMFQKXERI-QKAYSEPIDGVIHTGDTLEETAYFLNVEKTD 109
QY 117 DKPIVVGSMRPGTASADGMLNNAVAVASNDKSGKVLVTMDEIOSGRDVSSTIK 176
DB 110 KIPYILTGAMSSNEIGSDGLYNFISAIWTACSDSEYDKGVLVVWMDIHTARVYTKHT 169
QY 177 IKTEAPRS-AMGPIAGWVEGSKYWFRLPAKRHTVNSEFDIKQISLPQVDIAYSQVNTD 235
DB 170 TNVATFRTPPGTIGTAKERAFF---AKEVLPOEVCDVSSVKG--NVHVAKAYANGGE 223
QY 236 TAYKALQNGAKALIHAGTNGSVSRVFPALQELRNQVOIIRSRHQGGFYLRNAEOP 295
DB 224 RMFELANTPEIDGLVIELAGNLPETLPLALQKMLDNGIVLVSRCSNGI-----ADI 279
QY 296 DD-----KNDVVAHDINPQARILAMVAMTQDSKELQRIEMV 336
DB 280 YDVAGGVGLKMGVVFARGLNGPRARILIVGLNSKPNALKE-FLKH 328

RESULT 9

US-09-710-279-1912
/ Sequence 1912, Application US/09710279
/ Patent No. 6703492
/ GENERAL INFORMATION:
/ APPLICANT: KIMBERLY, WILLIAM JOHN
/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: PUJ480US

QY 128 PGTAMADGMLNLYNAVAASNDKSGVLTVMDEIOSGRDVSINIKTEAPKS-AW 186
 DB 127 SSNEIGSDGLYNPESAIRVASCEBALDKVLYVMDEIHSAHYVTHHTTNATPTPTPL 186
 QY 187 GFLGAVVEGKSYWFR--LPAGHTVNSBPDIOKISSLPQVDIAYSIGNVTDYAKALAN 244
 DB 187 GPICGVTAKRIIFLOBLLETKR-----LDISAVDG--TIPYKAYAGMOGDLLEIAHNT 238
 QY 245 GAKLIHAGTNGSVSRVVPALQELRKNGVOIIRSR-----QGGFVLRYA 292
 DB 239 KYDGLVTEALGAGNPPQALALEKLVKIPVLSRCPNGIAEPVYDEGGGKEL--- 295
 QY 293 EOPDDKNWVAVHDLNPOKARILAMVAMTKTODSKEL 329
 DB 296 ----EKMGITFCNSINSQKARVLLAVVYGLSGEHL 328

RESULT 14

US-09-543-681A-8255
 / Sequence 8255, Application US/09543681A
 / Patent No. 6605709
 / GENERAL INFORMATION:
 / APPLICANT: GARY BRETON
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 / FILE REFERENCE: 2709.1002-001
 / CURRENT APPLICATION NUMBER: US/09/543,681A
 / CURRENT FILING DATE: 2000-04-05
 / PRIOR APPLICATION NUMBER: US 60/128,706
 / PRIOR FILING DATE: 1999-04-09
 / NUMBER OF SEQ ID NOS: 8344
 / SEQ ID NO 8255
 / LENGTH: 347
 / TYPE: PRT
 / ORGANISM: Proteus mirabilis
 / US-09-543-681A-8255

Query Match 11.5%; Score 194; DB 4; Length 347;
 Best Local Similarity 24.2%; Pred. No. 8.3e-10;

Matches 75; Conservative 58; Mismatches 123; Indels 54; Gaps 13;

QY 11 NVVILATGTTAGAGASAAASATYQAAKGVVDKLIAGVPEL--ADLANVR-GEQVMQIAS 67
 DB 12 SIYVVYTGTT---GMQSHDHG-YIPVSGHLQRLAKMPEFRBEMPNFTIKELPLIDS 67
 QY 68 BITDDDLKASSVAELADSDVDGIVYTKGTDTLEETAYFLN-LVEKTDYPIVVGSM 126
 DB 68 SNITPDMQSIADDIS--NYCHYDGFVILAGDTMAFTASALSFMEGLAKPVIVTGSQ 125
 QY 127 RPTGMSADGMLNLYNAVAASNDKSGVLTVMDEIOSGRDVSINIKTEAPKSAM 186
 DB 126 IPELARSQGNLNLALYIAHHPT--NEVALPFNNNTLYRG--NRALKAHADGPN-- 178
 QY 187 GFLGAVVEGKSYWFR--LPAGHTVNSBPDIOKISSLPQVDIAYSIGNVTDYAKALAN 244
 DB 179 -----PASPNGAPLLEAGINIKTFINPLPKKGFIAHHTTPQIGVVTY 225
 QY 231 GNVTTAYYALANAKALI--HAGTNGSVSRVVPALQELRKNGVOIIRSR----- 282
 DB 226 PGLSBEVVNKKIIMQPKALILRSYGVNAPSHPALSTIREATRGVVVNLQICISGRV 285
 QY 283 QGGFVLRYA 292
 DB 286 NMGGYATGQA 295

RESULT 15

US-09-248-796A-17491
 / Sequence 17491, Application US/09248796A
 / Patent No. 674137
 / GENERAL INFORMATION:
 / APPLICANT: Keith Weinstein et al
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

QY 147 ASNKDSRGKGVLTVMDEIOSGRDVSINIKTEAPKS-AMGFLGAVVEGKSYWFRPLPAK 205
 DB 1 ASNRSGRGVLIANDRIGSGFYITKSNANSLDTFKSIGQSYGVNPFVNNRIYRPPPAK 60
 QY 206 -----RHTVNSER-----DIKQISSLPQVDIAYSIGNVTDYAKALAN-GAK 247
 DB 61 PGLTMFHINDLFSSYYVYALGNDI--IPCLPEVTIYAHQGLNNEIFKFIYNDLKSX 118
 QY 248 ALIHAGTNGSVSRVVPALQELRKNGVOIIRSRQGGF-----VLRYAE- 293
 DB 119 GIIILATMGAGSMDTNOYLSLNNPNYPIVYSKRSMDGWPPIGSLPKVLAVIKDAAD 178
 QY 294 ---QPDKN---DWVAVHDLNPOKARILAMVAMTKTODSKELQRTF 333
 DB 179 NKGPATKVPFVNAIAGTYLNPQKARILLQCLNGBMDSIKYVF 224

Query Match 11.4%; Score 191.5; DB 4; Length 227;
 Best Local Similarity 29.2%; Pred. No. 7.2e-10;

Matches 66; Conservative 39; Mismatches 80; Indels 41; Gaps 9;

Search completed: March 18, 2005, 21:34:11
 Job time : 45 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 21:26:43 ; Search time 141 Seconds
(without alignments)
787.673 Million cell updates/sec

Title: US-09-842-628-2

Perfect score: 1680
Sequence: 1 KEVENOQKLANVILATGTT.....MVAMTKQSKELQRIFFWEY 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications AA:
1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1680	100.0	336	9	US-09-842-628-2
2	758	45.1	375	14	US-10-406-025-2
3	717	42.7	364	10	US-09-882-227-84
4	356	21.2	320	17	US-10-472-928-4160
5	251	14.9	335	9	US-09-738-626-5846
6	127.5	7.6	548	15	US-10-369-493-513
7	114	6.8	544	15	US-10-282-122A-65021
8	113	6.7	544	15	US-10-282-122A-65715
9	113	6.7	723	15	US-10-282-122A-60463
10	113	6.7	1862	15	US-10-282-122A-49757
11	112	6.7	543	15	US-10-282-122A-64156
12	110.5	6.6	313	9	US-09-738-626-5967
13	110.5	6.6	330	16	US-10-781-014-504

14	110.5	6.6	545	15	US-10-282-122A-47225	Sequence 47225, A
15	110	6.5	540	15	US-10-369-493-181	Sequence 181, App
16	110	6.5	550	15	US-10-282-122A-50017	Sequence 50017, A
17	109.5	6.5	1574	16	US-10-695-499-179	Sequence 179, App
18	109.5	6.5	1978	16	US-10-695-499-60	Sequence 60, App
19	109.5	6.5	1981	9	US-09-928-457-38	Sequence 38, App
20	109.5	6.5	2015	14	US-10-066-551-1	Sequence 1, App
21	109.5	6.5	2015	15	US-10-282-122A-65772	Sequence 65772, A
22	109	6.5	423	16	US-10-437-965-147630	Sequence 147630, A
23	109	6.5	540	15	US-10-282-122A-74161	Sequence 74161, A
24	109	6.5	540	17	US-10-472-928-3964	Sequence 3964, App
25	109	6.5	641	14	US-10-267-311-51	Sequence 51, App
26	108.5	6.5	255	15	US-10-296-115-766	Sequence 766, App
27	108.5	6.5	540	15	US-10-282-122A-47926	Sequence 47926, A
28	108	6.4	613	15	US-10-282-122A-53710	Sequence 53710, A
29	107	6.4	542	15	US-10-369-493-10295	Sequence 10295, A
30	107	6.4	1057	15	US-10-282-122A-70305	Sequence 70305, A
31	107	6.4	1057	17	US-10-470-048B-152	Sequence 152, App
32	106.5	6.3	390	15	US-10-282-122A-67094	Sequence 67094, A
33	106.5	6.3	390	15	US-10-406-686A-32	Sequence 32, App
34	106.5	6.3	714	9	US-09-841-786-4	Sequence 4, App
35	106.5	6.3	3241	9	US-09-841-786-1	Sequence 1, App
36	106.5	6.3	3241	15	US-10-647-057-1	Sequence 1, App
37	106	6.3	543	15	US-10-282-122A-63595	Sequence 63595, A
38	106	6.3	4150	10	US-09-808-880-2	Sequence 2, App
39	105.5	6.3	923	9	US-09-969-362-1	Sequence 1, App
40	105.5	6.3	2457	15	US-10-282-122A-49854	Sequence 49854, A
41	105	6.2	449	15	US-10-282-122A-71468	Sequence 71468, A
42	105	6.2	2481	15	US-10-282-122A-43762	Sequence 43762, A
43	105	6.2	3413	17	US-10-042-665A-8	Sequence 8, App
44	104.5	6.2	6281	9	US-09-815-242-12996	Sequence 12996, A
45	104.5	6.2	10498	17	US-10-470-048B-440	Sequence 440, App

ALIGNMENTS

RESULT 1
US-09-842-628-2
Sequence 2, Application US/09842628
Patent No. US20020064862A1
GENERAL INFORMATION:
APPLICANT: ROBERTS, JOSEPH
APPLICANT: MACALLISTER, THOMAS W.
APPLICANT: SETHIRAMAN, NARARAJAN
APPLICANT: FREEMAN, ABIE G.
TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND ITS USE IN
FILE REFERENCE: 023032/0108
CURRENT APPLICATION NUMBER: US/09/842,628
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 08/050,482
PRIOR FILING DATE: 1995-04-25
PRIOR APPLICATION NUMBER: PCT/US92/10421
PRIOR FILING DATE: 1992-12-04
PRIOR APPLICATION NUMBER: DE P 4140003.8
PRIOR FILING DATE: 1991-12-04
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 336
TYPE: PRT
ORGANISM: Pseudomonas sp.
US-09-842-628-2

Query Match 100.0%; Score 1680; DB 9; Length 336;
Best local Similarity 100.0%; Pred. No. 7.9e-137;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEVENOQKLANVILATGTTAGAGSANSATYQAKGVNDLNGVPELDLNAVRE 60
DB 1 KEVENOQKLANVILATGTTAGAGSANSATYQAKGVNDLNGVPELDLNAVRE 60

QY 61 QWQJASESITNDLLKLAASVAELADSNVDGIVITHTGDTLEETAYELNVEKTDKPI 120
 DB 61 QWQJASESITNDLLKLAASVAELADSNVDGIVITHTGDTLEETAYELNVEKTDKPI 120
 QY 121 VVVGSRPCTASADGMLNNAVAASNDKSRGKGVLYTMNDEIQSOGDVSKSINIKTE 180
 DB 121 VVVGSRPCTASADGMLNNAVAASNDKSRGKGVLYTMNDEIQSOGDVSKSINIKTE 180
 QY 181 AFKSAWPGFMVVEGKSYWFLPAKHTVNSEFDIKQISLPQVDIAYSYGVTDTAYKA 240
 DB 181 AFKSAWPGFMVVEGKSYWFLPAKHTVNSEFDIKQISLPQVDIAYSYGVTDTAYKA 240
 QY 241 LAQNGAKALIHAGTNGSVSSRVVPAQLERKNGVQIIRSSROGGFVLRNAEQPDDKND 300
 DB 241 LAQNGAKALIHAGTNGSVSSRVVPAQLERKNGVQIIRSSROGGFVLRNAEQPDDKND 300
 QY 301 WYVADHNPQKARILAMVMTKQDSKELQRIFWEX 336
 DB 301 WYVADHNPQKARILAMVMTKQDSKELQRIFWEX 336

RESULT 2

US-10-406-025-2
 / Sequence 2, Application US/10406025
 / Publication No. US20030186380A1
 / GENERAL INFORMATION:
 / APPLICANT: No. US20030186380A1ozymes Biotech, Inc.
 / APPLICANT: Thomas, Michael D.
 / APPLICANT: Sloma, Alan
 / TITLE OF INVENTION: Method for producing secreted polypeptides having L-asparaginase
 / TITLE OF INVENTION: activity
 / FILE REFERENCE: 10289,200-US
 / CURRENT APPLICATION NUMBER: US/10/406,025
 / CURRENT FILING DATE: 2003-04-01
 / PRIOR APPLICATION NUMBER: US 60/369,192
 / PRIOR FILING DATE: 2002-04-01
 / NUMBER OF SEQ ID NOS: 8
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 2
 / LENGTH: 375
 / TYPE: PRT
 / ORGANISM: Bacillus subtilis
 / US-10-406-025-2

Query Match 45.1%; Score 758; DB 14; Length 375;
 Best Local Similarity 47.6%; Pred. No. 5,6e-57;
 Matches 158; Conservative 58; Mismatches 112; Indels 4; Gaps 3;

QY 6 QOKLANVILATGTTAGAGASANSATYQAAKVGVDKLIAGVPELADIANYRGEQWQI 65
 DB 47 KQDLPIRILATGGTTAGADQSKTSTTEYKAGVGVESLIEAVPEKDIANYSGEIQVNV 106
 QY 66 ABEISTNDLLKLAASVALADSNVDGIVITHTGDTLEETAYELNVEKTDKPIVVG 125
 DB 107 GSTINDKLLKLAARINHLASDDVDGIVITHTGDTLEETAYELNVEKTDKPIVVG 166
 QY 126 MRPGTASADGMLNNAVAASNDKSRGKGVLYTMNDEIQSOGDVSKSINIKTEAFKS- 184
 DB 167 MRPSTASISDGRSNLYNAKVAAPAKKGTLYVANDIABARYTKNTTTTIDPFSE 226
 QY 185 AWGLPMVVEGKSYWFLPAKHTVNSEFDIKQISLPQVDIAYSYGVTDTAYKALQN 244
 DB 227 EMGPGVTIAD-DIYFNEITRKHTKOTDSVSNLDELPOVDIITYGONGSYLFDAAVYA 285
 QY 245 GAKALIHAGTNGSVSSRVVPAQLERKNGVQIIRSSROGGFVLRNAEQPDDKND 304
 DB 286 GAKGIVFASGNGSISDAAEKADSAAKGVVAVSTRGNGVVPNOYAE--KDLAS 343
 QY 305 HDLNPQKARILAMVMTKQDSKELQRIFWEX 336
 DB 344 NSLNPQKARILAMVMTKQDSKELQRIFWEX 375

RESULT 3

US-09-882-227-84
 / Sequence 84, Application US/09882227
 / Publication No. US20030158396A1
 / GENERAL INFORMATION:
 / APPLICANT: Kleantous, Harold
 / APPLICANT: Al-Garawi, Amal
 / APPLICANT: Miller, Charles
 / APPLICANT: Tomb, Jean-Francois
 / APPLICANT: Oomen, Raymond P.
 / TITLE OF INVENTION: Identification of Polynucleotides
 / TITLE OF INVENTION: Encoding No. US20030158396A1el Helicobacter Polypeptides in the
 / FILE REFERENCE: 06132/047002
 / CURRENT APPLICATION NUMBER: US/09/882,227
 / CURRENT FILING DATE: 2001-06-15
 / PRIOR APPLICATION NUMBER: US 08/902,615
 / PRIOR FILING DATE: 1997-07-29
 / NUMBER OF SEQ ID NOS: 638
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 84
 / LENGTH: 364
 / TYPE: PRT
 / ORGANISM: Helicobacter pylori
 / US-09-882-227-84

Query Match 42.7%; Score 717; DB 10; Length 364;
 Best Local Similarity 44.6%; Pred. No. 1.9e-53;
 Matches 148; Conservative 75; Mismatches 103; Indels 6; Gaps 5;

QY 7 QOKLANVILATGTTAGAGASANSATYQAAKVGVDKLIAGVPELADIANYRGEQWQI 66
 DB 37 QNLPTIALATGTTAGAGAS-ASIGSYSGELGKELKAIPLNRLARIQEGEISNIG 95
 QY 67 SESITNDLLKLAASVAELADSNVDGIVITHTGDTLEETAYELNVEKTDKPIVVGSM 126
 DB 96 SODMNEBWFYKARQQLDSDRIQGVVITHTGDTLEESAFVNLVRSKTPVVLVGM 155
 QY 127 RPTGTASADGMLNNAVAASNDKSRGKGVLYTMNDEIQSOGDVSKSINIKTEAFKA- 185
 DB 156 RNAASISADGMLNNAVAASNDKSRGKGVLYTMNDEIQSOGDVSKSINIKTEAFKA- 215
 QY 186 WGPLMVEGKSYWFLPAKHTVNSEFDIKQISLPQVDIAYSYGVTDTAYKALQN 244
 DB 216 SGATISVYGGKTRRYMQPLRKHTTESSEFSLQKTPKPVIIYTHAGTIDPLFQASINS 275
 QY 245 GAKALIHAGTNGSVSSRVVPAQLERKNGVQIIRSSROGGFVLRNAEQPDDKND 304
 DB 276 HAKGVVINGVNGVNSAGFLKAMQASQMGVIVASSVNSGEI--TSGEIDDK-AFITS 332
 QY 305 HDLNPQKARILAMVMTKQDSKELQRIFWEX 336
 DB 333 DNLNPQKARVLLQALITKTNKKEKIQEMFEEX 364

RESULT 4

US-10-472-928-4160
 / Sequence 4160, Application US/10472928
 / Publication No. US20050020813A1
 / GENERAL INFORMATION:
 / APPLICANT: CHIRON SPA
 / APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
 / TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
 / FILE REFERENCE: P026926WO
 / CURRENT APPLICATION NUMBER: US/10/472,928
 / CURRENT FILING DATE: 2003-09-26
 / PRIOR APPLICATION NUMBER: GB-0107658.7
 / PRIOR FILING DATE: 2001-03-27
 / NUMBER OF SEQ ID NOS: 4979
 / SOFTWARE: SeqMan99, version 1.03
 / SEQ ID NO 4160
 / LENGTH: 320
 / TYPE: PRT

ORGANISM: Streptococcus pneumoniae
FEATURE:
OTHER INFORMATION: L-asparaginase, putative
OTHER INFORMATION: Cellular location: cytoplasm
OTHER INFORMATION: Similar to strain R6 sequence 15903854 (0.E+01)
US-10-472-928-4160

Query Match 21.2%; Score 356; DB 17; Length 320;
Best Local Similarity 31.6%; Pred. No. 2, 7e-22;
Matches 102; Conservative 51; Mismatches 134; Indels 36; Gaps 8;

QY 12 VVILATGGTGA-GAGASAAASATYQAAYGVDKLIAGVELADLANVRGEQVMOJASESI 70
DB 5 ILVLTGGTISQWADSGAVTSSDPMHNVSPLEGI-----QVHALDFNLPSPI 57
QY 71 TMDLLKASSVAELADSDVDGIVITHTGDTLEETAYFLNLVEKTDKPIYVVGSRPPT 130
DB 58 KPKHMLVLYQKKEED--NYDGVVITHTGDTLEETAYFLDMEVPHMIVLTGARSSN 115
QY 131 AMSADGMLNVAVAASNKDSRGKGLVTMNDIISGRDVSINIKTEAFKS-AMGPTL 189
DB 116 ELGSGGVNYSALNRVASDDRADKGVLVVMDIHAAYTKHTTNTVSTGTPTHGFL 175
QY 190 GNVGSKSYWFLPAKRTVNSEFDIKQISLPQVDIAYSYGNTDTAYKALAONGAKAL 249
DB 176 GLIMQELIYFTABPR---VRPDLDTIGL--VPIISAVGMTDELIDMLDLBHLDEL 229
QY 250 IHAGTNGSVSSRVVPALQELKNGVQITRRSR-----QQGFFVPMNEQPPD 297
DB 230 IIAQAGANGIPKRTQKESLLQKGPVALVSRFCFNGIAPVYAVGGGVOLQKA----- 284
QY 298 KNDVVAHDNPOKARIILAMVM 320
DB 285 --GVFPVKELNPKARKLILAL 305

RESULT 5

US-09-738-626-5846
Sequence 5846; Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASARO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5846
LENGTH: 325
TYPE: PRF
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5846

Query Match 14.9%; Score 251; DB 9; Length 325;
Best Local Similarity 26.2%; Pred. No. 3, 4e-13;
Matches 75; Conservative 53; Mismatches 120; Indels 38; Gaps 6;

QY 12 VVILATGGTGAAGASAAASATYQAAYGVDKLIAGVELADLANVRGE--QVMOJASES 69
DB 22 VAVITTTGGT--ACSDANNGHLLPTVS--GAD--LAPAPREFGAQIAFEIHINLSDSS 77
QY 70 ITNDLLKASSVAELADSDVDGIVITHTGDTLEETAYFLNLVEKTDKPIYVVGSRMG 129
DB 78 MTFEDLDSIIATVHVLYKLPDVVGAVVTHGDSMEBSAIAVDTPLDDPRPVLFTGAQKPF 137
QY 130 TMSADGMLNVAVAASNKDSRGKGLVTMNDIISGRDVSINIKTEAFKSAMGPTL 189
DB 138 DHPEDGPNNEPACLIASDPARGIIGALIVGHAIVPARGCVK----- 181
QY 190 GNVGSKSYWFLPAKRTVNSEFDIKQISLP-----QVDIAYSYGNTDTAYKALA 242
DB 182 -----MHTSDELAFTNGPEEPERPDALPAVKLADVSEIIPAYGATGATVEMAI 232
QY 243 ONGAKALIHAGTNGSVSSRVVPALQELKNGVQITRRSGQGGV 288
DB 233 AAGAGLVLEAMSGSVGSRMGDALGKALDAGIPVVMSTRVBRGEV 278

RESULT 6

US-10-369-493-513
Sequence 513; Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052) B
CURRENT APPLICATION NUMBER: US/10/369, 493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360, 039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 513
LENGTH: 548
TYPE: PRF
ORGANISM: Deinococcus radiodurans
US-10-369-493-513

Query Match 7.6%; Score 127.5; DB 15; Length 548;
Best Local Similarity 19.8%; Pred. No. 0.035;
Matches 94; Conservative 70; Mismatches 146; Indels 165; Gaps 21;

QY 1 KEVENOQKLVN-----VILATGGTGAAGASAA-----SATYQAK 38
DB 57 KEVELDCKLENTGAOLKEVASKTMDITGDTTITVAGQAIKGLRVNAAANPLAK 116
QY 39 VGVDDKLIAGVELADLANVRGEQVMOJASESITNDLLKASSVAELADSDVDGIVT 96
DB 117 RGIDRAVNAVAIEIKKLAVSVDSEAIKKNAGISANDE--TVGQELASMDKVGEGVIT 174
QY 97 ---THGDT-----LEETAYFL-----NL----- 112
DB 175 IEESGKGFTEVDVVEGMPDKGFINPYFTNDEKMAVLEDAVYIINEKKSINLXMDLPEV 234
QY 113 ---VEKTDKPIYVVGSRMGTSMA-----DGMNLNVAVAASNKDSRGKGLVTMND 163
DB 235 LERVAQGTGRPLILITADVEGELATLVNKLKGLTANI--AAVKAGFGGRKREMLRDI-A 291
QY 164 RIQSGRDVSKSINIKTEAFKSAMGPTGM-----VVEGSKSYWFLPAK 206
DB 292 AVTGERVVSSEDIGHLE-----NVGMEMLGRAARIRITKQETITVQKGRQAQIDARV 344
QY 207 HTVNSEFDI-----KOISLPQVDIAYSYGNTDTAYKAL-----AQNGAKAL 249
DB 345 NAIKELDLSTDSVAREKLOERTIAKLSGCVAVIRVGAATETELKEKGRHYEDALSTARS 404

Qy 250 IHAG--TANGSVSSRVVPALELRKN-----GVOI-----IRSSROGGFVL 289
Db 405 VEEGIVAGGGTTLRLVTPVRKAASLTGDEATGARILRLAEBPARQIAANAGEGSGV 464
Qy 290 RNABEPDK-----NDWVVAIDLNPOR-----ARILAMVMTX 322
Db 465 VNAVVGSDKARYGNATGEVYEDVMAAGIVDPARTALONAAISGALLITTE 519

RESULT 7

US-10-282-122A-65021
; Sequence 65021, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65021
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-282-122A-65021

Query Match 6.8%; Score 114; DB 15; Length 544;
Best Local Similarity 19.2%; Pred. No. 0.51;
Matches 80; Conservative 61; Mismatches 130; Indels 146; Gaps 19;

Qy 1 KEVENOQKLANV-----VILATGGINAGAS-----AANSATYQA-----K 38
Db 58 KEIKLKDPENNAQMVKEVASKTNDVAGDTTATVLAQSIIVAGMKVTVTGNMPTDK 117
Qy 39 VGVDDKILAG-VPELADLANV--RGEQVQIASESITNDLKLASSVAELADSDVDGIV 95
Db 118 RQIDAVNALVELKNIKAKPCPTSKKIAQVGSISANSDB--QVGALIAEMKREKRGV 175
Qy 96 ITHGDTLLE-----TAYPLNVER-----TDKPIVVGSMRPGTAMSADG 136

Db 176 TVEDEKSLLENLDVEGNOFDRGYSPIFYINDAEKOIAGLNDPVLDPDK--ISNIRD 232
Qy 137 MLNLVANAVALSN-----KDSRGKVLVTNNDIOSGRDVSXSINIKTEAF----- 182
Db 223 LLPVLEQVAKSRPLLIIAEVEGELATLYVNNI--RGLIKVAVAPGSDGRKML 289
Qy 183 -----KSANGPIGM-----VVEGKSIFYERLPK----- 205
Db 290 QDIALITGVVISEEVGLSEKATLDLGOAKRIEIGKENTTVIDGFDAAQIEARVAI 349
Qy 206 ---RHTVNSRD-----IKOISLPQVDIAYSGNTVDT--AYKALQN 244
Db 350 RQOJETATSDYDKELQERVAKLAGVAVITVGAATVEEMKDKORDVDAHATPAAYE 409
Qy 245 GAKALIHAGTNGSVSSRVVPALELRKN-----GVOIIRSSROGGFVLNABEP 295
Db 410 GVVA-----GGVVALLRARALNLTHTGNDDQDAGVOI-----VLRVAVSP 450

RESULT 8

US-10-282-122A-65715
; Sequence 65715, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65715
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-282-122A-65715

Query Match 6.7%; Score 113; DB 15; Length 544;
Best Local Similarity 18.9%; Pred. No. 0.62;
Matches 79; Conservative 62; Mismatches 130; Indels 146; Gaps 19;

```

: SEQ ID NO 60463
: LENGTH: 723
: TYPE: PR
: ORGANISM: Listeria monocytogenes
: US-10-282-122A-60463

Query Match          6.7%; Score 113; DB 15; Length 723;
Best Local Similarity 22.8%; Pred. No. 0.94;
Matches 74; Conservative 56; Mismatches 111; Indels 84; Gaps 17

QY 3 VENNQK---LANNVILATGGTIGAGAGSAANSATYQAKYGVDTLIGVPELADIANRG 59
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 378 LKNNRRLPLAKETKIALVGPILASSPPIIGGMNYYGHEKGIN-VENGLRNVFTVEYIS 436
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 60 EYVQWQJAS-----SITNDDLKLASSVAELADSDVDVGIVTGHGTLEETAYFL-N 111
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 437 TEYTFEPEDEKRVAKVAQNMDVVLL-----LQEKHEMGEGASLATIRLPEAQYELAK 491
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 112 LVKETDPIYV-VGSMP-----GRMSADGMLNY-----NAVNVASMKDR-- 153
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 492 FVQTLKRPVITLLENGRPLEVKEILAESSDALLELMPGTEAGRYADLLGASNPSEGLS 551
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 154 -----GKGVVITWMDIBIOSGRD-----VSKSINIKTEAFKSAWGPYGMVVEGKSY 198
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 552 MSPFGTGGQPVVYTNHLRTGRPQYPERKGERGVSHYLDIPHEPY-----PGY---GKSY 604
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 199 WFRLLPARKHTVNSBFDIKQISSLPQ-----VDIAVSYGVNTDTAVKALAQNGAKALIH 251
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 605 -----SFEELK-TSSLPEKELNLGEPLEHVEVTIKNISIDIAKEVYIQ-----VY 645
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 252 AGTNGSVSSRVVPALQELAKNGV 276
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 646 LDQVTASI-SRPVKEKAKFEKVALQ 669
   :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 10
US-10-282-122A-49757
: Sequence 49757, Application US/10282122A
: Publication No. US20040029129A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Liangsu
: APPLICANT: Zamudio, Carlos
: APPLICANT: Malone, Cheryl
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari
: APPLICANT: Zyskind, Judith
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John
: APPLICANT: Carr, Grant
: APPLICANT: Yamamoto, Robert
: APPLICANT: Forsyth, R.
: APPLICANT: Xu, H.
: TITLE OR INVENTION: Identification of Essential Genes in Microorganisms
: FILE REFERENCE: EITPA.034A
: CURRENT APPLICATION NUMBER: US/10/282,122A
: PRIOR FILING DATE: 2003-02-20
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/230,335
: PRIOR FILING DATE: 2000-09-06
: PRIOR APPLICATION NUMBER: 60/230,347
: PRIOR FILING DATE: 2000-09-09
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/267,636

```

;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 49757
;; LENGTH: 1862
;; TYPE: PRT
;; ORGANISM: Burkholderia fungorum
US-10-282-122A-49757

Query Match 6.7%; Score 113; DB 15; Length 1862;
Best Local Similarity 22.9%; Pred. No. 3.6;
Matches 78; Conservative 47; Mismatches 102; Indels 114; Gaps 16;

QY 15 LATGTTAGAGSA-ANSTYQAAYGVKLAGVPELADLANVGEQVMOJASISITND 73
DB 1044 LNAAGTIGGVGLAAGNATLTS---GRDPTDLG-----SLRGATIQGAGSGSTPA 1092
QY 74 DLKLASSVAELADSNVDGIVITHTGDTLEETAYFLNVEKDKPIYVVGSMRPGTAMS 133
DB 1093 DV--QAGSTIALYASNDVLTGTLAGSTVALTA-----GQDVVNSGTLQ---SSS 1138
QY 134 ADGMLNLYAVAVASNKSRG-GVL-----VTNDEIQSGRDV----- 171
DB 1139 DTNLALYSGVGVTSVNSGALNTVLAGTDINLGTTVALDPTLQSNADVAVTGSLSGQ 1198
QY 172 -----SSSINIKTEAFKSAMGPL--GMVVEKSYWFLPARHIVNSEFDIK 216
DB 1199 GNGFVTAGNIGAGSLAPAGSAVLNAGATISQGGILQGVN----- 1240
QY 217 QISSLPOVDIAYSIGNVDTAYKALAQ-----NGAKALIH--GTGNGSV----- 259
DB 1241 QVSAIGNAV---NNIESTSLALSAGTSGSGSLTVNGAALAAAITANGDVTVASA 1296
QY 260 -----SSRVPALQELRNQVOIIRSSHQGGFVLRNQ 294
DB 1297 GGLAAGTVGTALNDIN-----VAGATISNGDAVLNMQ 1331

RESULT 11
US-10-282-122A-64156
;; Sequence 64156, Application US/10282122A
;; Publication No. US20040029129A1
;; GENERAL INFORMATION:
;; APPLICANT: Wang, Ilangsu
;; APPLICANT: Zamudio, Carlos
;; APPLICANT: Malone, Cheryl
;; APPLICANT: Haseibeck, Robert
;; APPLICANT: Ohlsen, Karl
;; APPLICANT: Zyskind, Judith
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John
;; APPLICANT: Carr, Grant
;; APPLICANT: Yamamoto, Robert
;; APPLICANT: Forsyth, R.
;; APPLICANT: Xu, H.
;; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;; FILE REFERENCE: ELITRA.034A
;; CURRENT APPLICATION NUMBER: US/10/282,122A
;; CURRENT FILING DATE: 2003-02-20
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/230,335
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/230,347
;; PRIOR FILING DATE: 2000-09-09
;; PRIOR APPLICATION NUMBER: 60/242,578

;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/267,636
;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 64156
;; LENGTH: 543
;; TYPE: PRT
;; ORGANISM: Mycoplasma pneumoniae
US-10-282-122A-64156

Query Match 6.7%; Score 112; DB 15; Length 543;
Best Local Similarity 20.0%; Pred. No. 0.76;
Matches 83; Conservative 52; Mismatches 131; Indels 148; Gaps 18;

QY 1 KEVENQKLANV-----VILATGTTAGAGSANSATYQAAYGVKLAGVPELADL 54
DB 57 KEIELDPLENIGAKVISVAAVSTNDIAGDGTTRTTLAQETNRGVAHVNNG---ANP 112
QY 55 ANVR-----GEQVMOJASISITNDLKLASSVAELADSNVD 91
DB 113 VNVRGIDASQILITELDKRSKKINTMBEIQVALISSGKEIGKLAQAMALVGN-- 170
QY 92 DGIYHTGDTLEET-----AYFLNVEK---TDKPIVVV----- 123
DB 171 -GVITTDPAKINTITLETGIEFEKGYASPYMSDQKMEVLDQPIVLVSAMKINTIK 229
QY 124 -----GSMRPG-----TAMSDGMLNLYAVAVASNK--DSRGKGV-- 158
DB 230 EILPLBSGMEGNPLIVADDPAEVVTTLAVNGLRGTINVAACKMEYGERQALLED 289
QY 159 -----VTNDEIQSG-RDYSKSIKITEAFKSAMGPLGMVVEKSYWFLPARHIVN 210
DB 290 LAISGTLAVNNELGGFKDV--TVNHLGEARV-----QVAKETTVIGSGSKETIQ 341
QY 211 SEFDI-----KOISLPOVDIAYSIGNVDTAYKAL-----AONGA 246
DB 342 KHLDLNGLKQTEKYDTDLKERIAHLSQGVAVVRVGATELAQKELKRIEDALNST 401
QY 247 KALIHAG--TGNGSSSRVVPALQELR-----KN-----GVOIIRSS 281
DB 402 KAAVEGIIISGGIALNLNVSTIINDSKLADYKAKETSANLKEILVGEIYAKS 455

RESULT 12
US-09-738-626-5967
;; Sequence 5967, Application US/09738626
;; Publication No. US20020197605A1
;; GENERAL INFORMATION:
;; APPLICANT: NAKAGAWA, SATOSHI
;; APPLICANT: MIZOGUCHI, HIROSHI
;; APPLICANT: ANDO, SEIKO
;; APPLICANT: HAYASHI, MIKIRO
;; APPLICANT: OCHIAI, KEIKO
;; APPLICANT: YOKOI, HARUHIKO
;; APPLICANT: TATEISHI, NAOKO
;; APPLICANT: SENOH, AKIHIRO
;; APPLICANT: IKEDA, MASATO
;; APPLICANT: OKAZI, AKIO
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-125
;; CURRENT APPLICATION NUMBER: US/09/738,626
;; CURRENT FILING DATE: 2000-12-18
;; PRIOR APPLICATION NUMBER: JP 99/377484
;; PRIOR FILING DATE: 1999-12-16
;; PRIOR APPLICATION NUMBER: JP 00/159162

PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5967
LENGTH: 313
TYPE: PRF
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5967

Query Match 6.6%; Score 110.5; DB 9; Length 313;
Best Local Similarity 22.0%; Pred. No. 0.46;
Matches 69; Conservative 53; Mismatches 120; Indels 71; Gaps 15;

QY 15 LATGTGAGAGSAAASATYQAAYGVVDL-----AGPELADLAN 56
DB 3 LASCSSDSDSSTSDAGDSYRVGINQLVGHAPALDAATTGPKAFBAGVDVTFDEQN 62
QY 57 VAGEQ-----VMOJASESITNDLLKLAASVAELADSNVD-GIVTHTDTLSEYAF 109
DB 63 ANGEQGTALTISQAFASDWL--DLVLAATPPAQATQNTIDIPVLFATVD-----AVS 115
QY 110 LNLVEKTKPIVVGSMRPGTAMSA--DQMLNLYNAVAASNKDSRGKGLVTMNDLQIS 167
DB 116 AELVDSNEAP-----GGNTVGTSDIAPISQQLLELLQGLV---PDKSGIGI-VYASGEVNS 166
QY 168 GBDVKSINIKTEAFKASANGPLGMYVEGKSYWFLPAKHNTNSEFDIQ-ISSLPEVDI 226
DB 167 -----QVQVDEYTKAEPPLGLSV-----NTQYVTVNRIQAAVEALGDVDV 207
QY 227 AY-SYGVNTDTYVKLAONGAKALHA-GTNGSVSRRVPAQLGRKNGVQIIRSSROQ 284
DB 208 IYVPTDNMVSIGISLVQVABEQKLPVIGASGIVEGALATL-----GIDYTELGKQT 261
QY 285 GGFVLRNAEQPDD 297
DB 262 GEMALRIIQDGED 274

RESULT 13
US-10-781-014-504
Sequence 504, Application US/10781014
Publication No. US20040180408A1

GENERAL INFORMATION:
APPLICANT: Pompeju, Markus
APPLICANT: Krogger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
FILE REFERENCE: BGI-126CPN
CURRENT APPLICATION NUMBER: US/10/781,014
PRIOR FILING DATE: 2004-02-17
PRIOR APPLICATION NUMBER: US 09/602,740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/143,208
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/151,572
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19931412.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931413.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931424.1

PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931428.4
PRIOR FILING DATE: 1999-07-08
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 784
SEQ ID NO 504
LENGTH: 330
TYPE: PRF
ORGANISM: Corynebacterium glutamicum
US-10-781-014-504

Query Match 6.6%; Score 110.5; DB 16; Length 330;
Best Local Similarity 22.0%; Pred. No. 0.5;
Matches 69; Conservative 53; Mismatches 120; Indels 71; Gaps 15;

QY 15 LATGTGAGAGSAAASATYQAAYGVVDL-----AGPELADLAN 56
DB 20 LASCSSDSDSSTSDAGDSYRVGINQLVGHAPALDAATTGPKAFBAGVDVTFDEQN 79
QY 57 VAGEQ-----VMOJASESITNDLLKLAASVAELADSNVD-GIVTHTDTLSEYAF 109
DB 80 ANGEQGTALTISQAFASDWL--DLVLAATPPAQATQNTIDIPVLFATVD-----AVS 132
QY 110 LNLVEKTKPIVVGSMRPGTAMSA--DQMLNLYNAVAASNKDSRGKGLVTMNDLQIS 167
DB 133 AELVDSNEAP-----GGNTVGTSDIAPISQQLLELLQGLV---PDKSGIGI-VYASGEVNS 183
QY 168 GBDVKSINIKTEAFKASANGPLGMYVEGKSYWFLPAKHNTNSEFDIQ-ISSLPEVDI 226
DB 184 -----QVQVDEYTKAEPPLGLSV-----NTQYVTVNRIQAAVEALGDVDV 224
QY 227 AY-SYGVNTDTYVKLAONGAKALHA-GTNGSVSRRVPAQLGRKNGVQIIRSSROQ 284
DB 225 IYVPTDNMVSIGISLVQVABEQKLPVIGASGIVEGALATL-----GIDYTELGKQT 278
QY 285 GGFVLRNAEQPDD 297
DB 279 GEMALRIIQDGED 291

RESULT 14
US-10-282-122A-47225
Sequence 47225, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 47225
 ; LENGTH: 545
 ; TYPE: PRT
 ; ORGANISM: Borrelia burgdorferi
 ; US-10-282-122A-47225

Query Match 6.6%; Score 110.5; DB 15; Length 545;
 Best Local Similarity 18.0%; Pred. No. 1;
 Matches 73; Conservative 68; Mismatches 131; Indels 133; Gaps 17;

QY 1 KEVENOQKLANV-----VILATGTTAGASAAANSATTGAQKGVYDKLIAGVPELA- 52
 DB 57 REIELENPENNGAQLKEVALKT-NDVAGDGTATATVLAIAEBGLKNVSSGINPIGI 115
 QY 53 -----DLANVRGEQVWQIASISITNDLKLAS-----SVAELADSDVDGIYI 96
 DB 116 KKGIDHAVVLAIAEKIRQSAKKTITKEEIAQVASISANDSYIGEKIAEMDKVGDGYIT 175
 QY 97 THGIDTLEET-----AYF-----LNL 112
 DB 176 VEESKTFDTTISYVGMPDRGYLSPFSTNKENSVDNDAFILIYEKLISSIKELLPV 235
 QY 113 VKE-----TDKPIYVVGSMRPGTMSADGMLNTYNAVAVASNKD-----SRGKGLVTMNDBI 165
 DB 236 LEKVTGTFPLIILADIEGDALALVNSVKGALVCAIKSPGCGDRKAKMLEDI-AVL 294
 QY 166 QSGRDVSKSINKTEAFKSAWGPLGVVE--GKSYWFLPAKRTV---NSEPDIKOTS 219
 DB 295 TGGVLAISEELGLTLEVE-----IKQLGQAKTIKVDKONTTILINTGNE-QIKERS 344
 QY 220 SLPOVDI-----AYSQNTDTIYKAL-----AQNGAKAL 249
 DB 345 ELIKQIEDSTSEYDEKLOERLAKLVGVAVINGAIVTEVELEKKEHREVDALSATRAA 404
 QY 250 IHAGT--GNGSVSRVVPALQEL-----RKNGQIIRSSROQ 284
 DB 405 VEEGVVPGGGSTLIEVAMTLDITITSKLSYERKQGEIYKSLER 449

RESULT 15
 US-10-369-493-181
 ; Sequence 181, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 181
 ; LENGTH: 540
 ; TYPE: PRT
 ; ORGANISM: Xenorhabdus nematophilus
 ; US-10-369-493-181

Query Match 6.5%; Score 110; DB 15; Length 540;
 Best Local Similarity 18.8%; Pred. No. 1.1;
 Matches 85; Conservative 73; Mismatches 141; Indels 154; Gaps 20;

QY 1 KEVENOQKLANVILATGTTAGASAAANSATTGAQKGVYDKLIAGVPELA- 37
 DB 58 REIELENPENNGAQLKEVALKT-NDVAGDGTATATVLAIAEBGLKNVSSGINPIGI 112
 QY 38 -----KVGVDK-LIAGVPELADLANVRGEQ--VMQIASISITNDLKLASSVAELADSD 90
 DB 113 PMDLKRGIDKAVIAAVERELKLSVPCSDSTIAIAGVGTISANSDE--TVGKLIABEMDKYG 170
 QY 91 VDGIV-ITHGIDTLEET-----TAYPLNT----- 112
 DB 171 KEGVITVEGTGLDELDVVEGMPDRGYLSPFINKPEAGSILENPFYILLVDKISNI 230
 QY 113 -----VEKTDKPIYVVGSMRPGTMSADGMLNTYNAVAVASNKD-----DSRGKGLV 159
 DB 231 RELPLVEBGVAKASPLVILAEVGEALATLVNNMGIVKVAVAKAPGDRKAKMLQ 290
 QY 160 TMNDEIOSGRDVSXKINKTEAFKSAWGPLG-----HVEGKSYWFLPAK-- 205
 DB 291 DI-ATLNGTVIISIEIGLEL--KATLEDLQAKRIYINKOTTTIIDVGEMAAIAAVT 347
 QY 206 --RHTV---NSEPD-----IKQISLPOVDIAYSQNVDTT--AYKALA 242
 DB 348 QIROQIEESTSDYREKLOERLAKLVGVAVINGAIVTEVELEKKEHREVDALSATRAA 407
 QY 243 QNG-----AKALIHGTGNG--SVSRVVPALQELRKNGVGTIIRSSROQGFV 288
 DB 408 EGVVAGGVALVRAAAIISALTGDNDQNVGIRV--ALRAEAPMRQIVENAGEBPSVY 465
 QY 289 LRNAE-----QPDCKNDWVAHDLPK 311
 DB 466 VNNVKAQSGNHTGYNAATEEYGDMEIMGLDPTK 498

Search completed: March 18, 2005, 21:36:39
 Job time : 144 secs

A28063 (EC 3.5.1.38) - Acinetobacter calcoaceticus
N:Alternate names: glutaminase-asparaginase
C:Species: Acinetobacter calcoaceticus
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A28063
R:Tanaka, S.; Robinson, B.A.; Appella, E.; Miller, M.; Ammon, H.L.; Roberts, J.; Weber, J. Biol. Chem. 263, 8583-8591, 1988
A:Title: Structures of amidohydrolases. Amino acid sequence of a glutaminase-asparaginase hysantheml.
A:Reference number: A28063; MUID:88243706; PMID:3379033
A:Accession: A28063
A:Molecule type: protein
A:Residues: 1-331 <TAN>
A:Cross-references: UNIPROT:P10172
A>Note: the source is designated as Acinetobacter glutaminasificans
C:Superfamily: asparaginase
C:Keywords: hydrolase

Query Match 61.8%; Score 1037.5; DB 1; Length 331;
Best Local Similarity 62.0%; Pred. No. 2,2e-63;
Matches 204; Conservative 52; Mismatches 70; Indels 3; Gaps 2;

QY 11 NVVILATGTTAGAGAAASATYQAKYGVDTLIGVPELADLANVGEQVMQIASBI 70
DB 3 NVVIVATGTTAGAGASTNSATYSNAKYPVDALIKAVQVNDLANITGICALQVASEBI 62
QY 71 TNDLILKLASSVAELADSDVDGIVTHTGDTLLEAFYPLNVEKTDKPIVVGSRPPT 130
DB 63 TDKELSLARQVNDLVKPSVGVVITHTGDTMEETAFPLNVHTDKPIVVGSRPST 122
QY 131 AMSADGMLNLVNAVAVASNKDSRGKGLVTMNDLISGSDVSKSINIKTEAFKSAWGPGL 190
DB 123 ALSADGPTLVSAVALASSENKKNQVWVLMDSIFPAADVTGKINIHHTAFVSGWALG 182
QY 131 MYVEGSKYWFRLPAKHTVNSBFDIKQI--SLPQVDIAVSYGNTDTAYKALAONGAK 248
DB 183 TLVEKGPYWFRRSVKGTITNSSEFNTEKIQGDALPGQIYVGSDDNMPDAPYQAPAKAGVCA 242
QY 249 LIHAGTNGSVSSRVVPAQLGR-KNGVQIIRSSRQGGFVLRNAEQPDKNQVVAHDL 307
DB 243 ITHAGTNGSMANVLPVPRKLDHDEGLQIVRSSRVAGFVLRNAEQPDKIGWIAHDL 302
QY 308 NPQKRIILAMVAMTKTQDSKELQRIFWEX 336
DB 303 NPQKRIILMALATLTNTDAKEIQNMFNXY 331

RESULT 3
D85953
periplasmic L-asparaginase II [imported] - Escherichia coli (strain O157:H7, substrain H
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85953
R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
R:iller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimailanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85953
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <STO>
A:Cross-references: UNIPROT:O8XC02; GB:AB005174; NID:G12517506; PIDN:AA058088.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ansb
C:Superfamily: asparaginase

Query Match 46.8%; Score 786; DB 2; Length 348;
Best Local Similarity 48.3%; Pred. No. 3.2e-46;
Matches 159; Conservative 59; Mismatches 107; Indels 4; Gaps 3;

QY 9 LANVILATGTTAGAGAAASATYQAKYGVDTLIGVPELADLANVGEQVMQIASB 68
DB 23 LPNITILATGTTAGAGGDSATKS-NYTAGKGVENVAVVQLQDIANVKGQVNNISQ 81
QY 69 SITNDLILKLASSVAELADSDVDGIVTHTGDTLLEAFYPLNVEKTDKPIVVGSRP 128
DB 82 DMDNDVMTLLAKKIN--TDCCKTQDFVITHTGDTMEETAFYPLDVLVKKDCKPVMWGANRP 139
QY 129 GTAMSDGMLNLVNAVAVASNKDSRGKGLVTMNDLISGSDVSKSINIKTEAFKSA-WG 187
DB 140 STMSADGPTLVNANVTAADKASANRGVLVMDVTLDGSDVTKNTTDTVATFKSVNYG 199
QY 188 PLGMYVEGSKYWFRLPAKHTVNSBFDIKQISLPQVDIAVSYGNTDTAYKALAONGAK 247
DB 200 PLGYIHNGKIDYQRTPAKHTSDPFPVSKINELPKVGI VVNYNAAADLPKALVDAGVD 259
QY 248 ALIHAGTNGSVSSRVVPAQLGRNGVQIIRSSRQGGFVLRNAEQPDKNQVVAHDL 307
DB 260 GIVSAGVNGNVLKYSVFPTTLTAANKGTAVVRSRVPFGATTQDAKVDKYGFIASGTL 319
QY 308 NPQKRIILAMVAMTKTQDSKELQRIFWEX 336
DB 320 NPQKRIILQALATVOTDPOQIQOIFNOY 348

RESULT 4
A88108
periplasmic L-asparaginase II [imported] - Escherichia coli (strain O157:H7, substrain R)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: A88108
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasaewara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A:Reference number: A89629; MUID:21156231; PMID:11258756
A:Accession: A88108
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <HAY>
A:Cross-references: UNIPROT:O8XC02; GB:BA000007; PIDN:BA037256.1; PID:G13363305; GSPDB:G
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC03833
C:Superfamily: asparaginase

Query Match 46.8%; Score 786; DB 2; Length 348;
Best Local Similarity 48.3%; Pred. No. 3.2e-46;
Matches 159; Conservative 59; Mismatches 107; Indels 4; Gaps 3;

QY 9 LANVILATGTTAGAGAAASATYQAKYGVDTLIGVPELADLANVGEQVMQIASB 68
DB 23 LPNITILATGTTAGAGGDSATKS-NYTAGKGVENVAVVQLQDIANVKGQVNNISQ 81
QY 69 SITNDLILKLASSVAELADSDVDGIVTHTGDTLLEAFYPLNVEKTDKPIVVGSRP 128
DB 82 DMDNDVMTLLAKKIN--TDCCKTQDFVITHTGDTMEETAFYPLDVLVKKDCKPVMWGANRP 139
QY 129 GTAMSDGMLNLVNAVAVASNKDSRGKGLVTMNDLISGSDVSKSINIKTEAFKSA-WG 187
DB 140 STMSADGPTLVNANVTAADKASANRGVLVMDVTLDGSDVTKNTTDTVATFKSVNYG 199
QY 188 PLGMYVEGSKYWFRLPAKHTVNSBFDIKQISLPQVDIAVSYGNTDTAYKALAONGAK 247
DB 200 PLGYIHNGKIDYQRTPAKHTSDPFPVSKINELPKVGI VVNYNAAADLPKALVDAGVD 259
QY 248 ALIHAGTNGSVSSRVVPAQLGRNGVQIIRSSRQGGFVLRNAEQPDKNQVVAHDL 307
DB 260 GIVSAGVNGNVLKYSVFPTTLTAANKGTAVVRSRVPFGATTQDAKVDKYGFIASGTL 319
QY 308 NPQKRIILAMVAMTKTQDSKELQRIFWEX 336
DB 320 NPQKRIILQALATVOTDPOQIQOIFNOY 348

RESULT 5

asparaginase (EC 3.5.1.1) II precursor [validated] - *Escherichia coli* (strain K-12)

C/Species: *Escherichia coli*
C/Date: 24-Apr-1984 #sequence revision 31-Dec-1990 #text_change 09-Jul-2004

A/Accession: A35132; J03001; A01000; D65081
R/Jerninger, M.P.; Beacham, I.R.

J/Bacteriol. 172, 1491-1498, 1990

A/Title: Analysis of the *Escherichia coli* gene encoding L-asparaginase II, anSB, and the

A/Reference number: A35132; PMID:90170867; PMID:2407723

A/Accession: A35132

A/Molecule type: DNA

A/Residues: 1-348 <JEN>

A/Cross-references: UNIPROT:P00805; EMBL:X52540

R/Bonthon, D.T.

Gene 91, 101-105, 1990

A/Title: The primary structure of *Escherichia coli* K-12: cloning, mapping and sequencing of

A/Reference number: J03001; PMID:90382683; PMID:2144836

A/Accession: J03001

A/Molecule type: DNA

A/Residues: 1-348 <BON>

A/Cross-references: GB:M34234; NID:G145276; PIDN:AAA23445.1; PID:G145277

A/Experimental source: strain K12 JM108

R/Maita, T.; Matsuda, G.

Hoppe-Seyler's Z. Physiol. Chem. 361, 105-117, 1980

A/Title: The primary structure of L-asparaginase from *Escherichia coli*.

A/Reference number: A01000; PMID:80135739; PMID:676894

A/Accession: A01000

A/Molecule type: protein

A/Residues: 23-48 'A', 50-85 'D', 87-131, 133-155, 157-170, 172-205, 'D', 207-267, 'D', 269-273, 'R' Peterson, R.G.; Richards, P.F.; Handschumacher, R.E.

J. Biol. Chem. 252, 2072-2076, 1977

A/Title: Structure of peptide from active site region of *Escherichia coli* L-asparaginase

A/Reference number: A32655; PMID:77140944; PMID:321449

A/Accession: A32655

A/Contents: annotation; active site

R/Greenquist, A.C.; Wriston Jr., J.C.

Arch. Biochem. Biophys. 152, 280-286, 1972

A/Title: Chemical evidence for identical subunits in L-asparaginase from *Escherichia coli*

A/Reference number: A37451; PMID:73007901; PMID:4561256

A/Accession: A37451

A/Contents: annotation

A/Note: The cysteine residues were quantitated and shown to form intrachain bonds

R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

Science 277, 1453-1462, 1997

A/Title: The complete genome sequence of *Escherichia coli* K-12.

A/Reference number: A64720; PMID:97426617; PMID:9278503

A/Accession: D65081

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-348 <BLAT>

A/Cross-references: GB:AE000378; GB:U00096; NID:G1789319; PIDN:AACT5994.1; PID:G1789327;

A/Experimental source: strain K-12, substrain MGL655

C/Genetics:

A/Genes: anSB

A/Map position: 64 min

C/Complex: homotetramer

C/Function: EC 3.5.1.1 [validated, PMID:90170867]; catalyzes the hydrolysis of aspara

A/Note: has a higher affinity for asparagine than asparaginase I (PIR:XDCI)

A/Note: positively regulated by CAMP receptor protein (CRP) (PIR:QRECC) and by PNR prote

QY 9 LANVILATGGTITAGAGSAAASATYQAKVGVDKLIAGVPEIADLANVGRQWQIASSEIT 68

DB 23 LPNITITLATGGTITAGGGSATKS -HYTGKGVENLVNVAVPPQKDIQIANVKGQVNIISGQ 81

QY 69 SITNDLKLASSVAELADSDVDGIVITHTGDTLEETAYFLNLYEKTDPYVVGSGMRP 128

DB 82 DNDNDVWVTLAKKIN -TDGCTKDPFVITHTGDTLEETAYFLNLYEKTDPYVVGSGMRP 139

QY 129 GTMSADGMLNLYNVAVANSNDGKGVLYMNDIEGSRVGSINIKTEAFPSA -WG 187

DB 140 STMSADGFMNLYNVAVNTADASANRGLVVMNDIVLDGRVYTKNTTVDATFNSVNTG 199

QY 188 PLGMVVEGKSYWFRLLPAKHVTYNSERFDIKQISLPPQVDIAYSQVNTDTAYALQNGAK 247

DB 200 PLGYIHNGKIDYQRPARKHTSDTPEVSKNELPKVGLVYVYANASDLPAKALVDAGYD 259

QY 248 ALIHAGTNGSVSRVPAFLQRLKNGVOIIRSSQGGFVLNABQPDKNDWVVAHD 307

DB 260 GIVSAGVNGNLYKSVFDTLTAAKTGAVVRSRVPTGATYQDAVDAKGFVASGTL 319

QY 308 NPQKARILLAMVMTKQDSKEIQRIFWEY 336

DB 320 NPQKARVLLQALTYQKQDQIQLFNQY 348

RESULT 6

S74205

asparaginase (EC 3.5.1.1) - *Molinitella succinogenes*

C/Species: *Molinitella succinogenes*

C/Date: 04-Dec-1997 #sequence revision 12-Dec-1997 #text_change 09-Jul-2004

A/Accession: S74205; S52788; S57754

R/Rubrowek, J.; Palm, G.J.; Gilliland, G.L.; Derst, C.; Roehm, K.H.; Wlodawer, A.

Bur. J. Biochem. 241, 201-207, 1996

A/Title: Crystal structure and amino acid sequence of *Molinitella succinogenes* L-asparaginase

A/Reference number: S74204; PMID:97054610; PMID:8898907

A/Accession: S74205

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-330 <LUB>

A/Cross-references: UNIPROT:P50286; EMBL:X89215; NID:G885917; PIDN:CAA61503.1; PID:G88591

A/Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1995

R/Derst, C.; Roehm, K.H.

A/Description: Cloning and sequencing of L-asparaginase from *Molinitella succinogenes*.

A/Reference number: S52788

A/Accession: S52788

A/Molecule type: DNA

A/Residues: 1-330 <DER>

A/Cross-references: EMBL:X83689; NID:G758651; PIDN:CAA58658.1; PID:G758652

C/Genetics:

A/Genes: anSA

C/Superfamily: asparaginase

C/Keywords: homotetramer; hydrolase

F/14,93,94,166/Active site: Thr, Asp, Lys #status predicted

Query Match 45.5%; Score 765; DB 2; Length 330;

Best Local Similarity 49.4%; Pred. No. 8, 1e-45;

Matches 161; Conservative 51; Mismatches 112; Indels 2; Gaps 2;

QY 12 VILATGGTITAGAGSAAASATYQAKVGVDKLIAGVPEIADLANVGRQWQIASSEIT 71

DB 6 VILATGGTITAGGGSATKS -HYTGKGVENLVNVAVPPQKDIQIANVKGQVNIISGQ 64

QY 72 NDLLKLASSVAELADSDVDGIVITHTGDTLEETAYFLNLYEKTDPYVVGSGMRP 131

DB 65 GKWTLLAKRVNELLQKTEBAVITHTGDTLEETAYFLNLYEKTDPYVVGSGMRP 124

QY 132 MSADGMLNLYNVAVANSNDGKGVLYMNDIEGSRVGSINIKTEAFPSA -WGPG 190

DB 125 MSADGFMNLYNVAVNTADASANRGLVVMNDIVLDGRVYTKNTTVDATFNSVNTG 184

QY 191 MVVEGKSYWFRLLPAKHVTYNSERFDIKQISLPPQVDIAYSQVNTDTAYALQNGAKAL 250

Db 165 TTYGKVEYFQSVRPHLASEPDISKIBELPRVDILVAHPDDTDLVNAALQAGKII 244
 Qy 251 HAGTNGSVSRVPAALQELRKNGVQIIRSSRQGGFVLRNAEQPDDKNDVVAADLNPQ 310
 Db 245 HAGNGNGNFPILQALRBAKASGVVARSVSGSTTQBAFVDDKLGAFATSIANPQ 304
 Qy 311 KARIILAMVAMTKTQDSKEIQRIFWEX 336
 Db 305 KARVILMLALTKTSDREAIQKIFSTY 330

RESULT 7

Asparaginase (EC 3.5.1.1) - *Salmonella enterica* subsp. *enterica* serovar Typh1 (strain CT A80879)
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh1
 A:Note: this species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AB0879
 R:Parthill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, R.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typh1
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AB0879
 A:Status: preliminary
 A:Statues: preliminary
 A:Molecule type: DNA
 A:Residues: 1-348 <PAR>
 A:Cross-references: GB:AL513382; PIND:CAD02930.1; PID:g16504183; GSPDB:GN00176
 C:Gene: STY3259
 C:Superfamily: asparaginase
 C:Keyworte: hydrolase

Query Match 45.3%; Score 761; DB 2; Length 348;
 Best Local Similarity 47.1%; Pred. No. 1.6e-44;
 Matches 151; Conservative 60; Mismatches 110; Indels 4; Gaps 3;

Qy 9 LANVYILATGTTAGAGSANSATYQAKGVVDKLIAGVPELADLANVGRGVQIASE 68
 Db 23 LPNITLITLGTITAGGDSATKS-NTYACKGVEVNLVDVPOKDIAYVKGQVANNISQ 81
 Qy 69 SITNDLLKLASSVAELADSDVDGIVITHGTDLBETAYFLNVEKTDKPIVVGSMRP 128
 Db 82 DWNDEWVLTIAKKINTBCDS--TDGFTVTHGTDMETKAYFLDLTVCKNKPVLVGARRP 139
 Qy 129 GRVMSADGMLNLYNAVAVASNDKSRGKGLVYTMNDEIQSGRVSISINIKTAFKSA-WG 187
 Db 140 STSMGADGPFNLVYNAVVTADQSANRGLVYVMDTVMDGRDVTIKNTTDDVATFFAVVYG 199
 Qy 168 PLGMVVEGKSYFRLPAKHHTVNSSEFDIQISLPQVDIAYSGVNTDTAYKALQNGAK 247
 Db 200 PEGYIHNGKIDYQRTPERKHTTSTPFDVSKTALRKEVGIYVYNAASDLPAKALVDAGTD 259
 Qy 248 ALIHGNGSVSRVPAALQELRKNGVQIIRSSRQGGFVLRNAEQPDDKNDVVAADL 307
 Db 260 GIVSAGVNGNLYKYVPTDLATRAHNGTVVAVSSRPVATATTQDAEVDKAGVFAASGL 319
 Qy 308 NPQKARILAMVAMTKTQDSKEIQRIFWEX 336
 Db 320 NPQKARVILQALTKTQDKPKVIOQIFEDP 349

RESULT 8

Asparaginase (EC 3.5.1.1) II precursor - *Haemophilus influenzae* (strain Rd KW20)
 C:Species: *Haemophilus influenzae*
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
 C:Accession: A64090
 R:Plieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.; Goeyens, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrman, J.L.; Geoghegan, N.S.M.

Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: A64090
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-349 <TIGR>
 A:Cross-references: UNIPROT:P43843; GB:U32758; GB:L42023; NID:g1573747; PIND:AA022403.1;
 C:Superfamily: asparaginase
 C:Keywords: extracellular protein; hydrolase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-349/Product: asparaginase #status predicted <MAT>
 F:100-128/Distilled bonds: #status predicted

Query Match 45.1%; Score 758; DB 2; Length 349;
 Best Local Similarity 45.1%; Pred. No. 2.6e-44;
 Matches 151; Conservative 74; Mismatches 106; Indels 4; Gaps 3;

Qy 3 VENQOKLANVYILATGTTAGAGSANSATYQAKGVVDKLIAGVPELADLANVGRGV 62
 Db 18 IANAADLPNITLITLGTITAGGDSATKS-YKAGQLSIDTLIBAVPEKNTIANIKGRQI 76
 Qy 63 MOIASSEITNDLLKLASSVAELADSDVDGIVITHGTDLBETAYFLNVEKTDKPIYV 122
 Db 77 VKIGSQDNDDEVMLKLAIAIN--AQCKSTDPFTVTHGTDMETKAYFLDLTVCKEKPVL 134
 Qy 123 VGSNRPGTAMSADGMLNLYNAVAVASNDKSRGKGLVYTMNDEIQSGRVSISINIKTAF 182
 Db 135 VGANRPATEKSNADGMLNLYNAVVAADKSSRGVLYVMDTVMDGRDVTIKNTTDDVAT 194
 Qy 163 KSA-WGPIGMVVEGKSYFRLPAKHHTVNSSEFDIQISLPQVDIAYSGVNTDTAYKAL 241
 Db 195 HSPNYGSLGYIHNSKVDYERSPEKHTINTPENVKLDLSPVGIYVYNAVPEPLNAL 254
 Qy 242 AONGAKALIHGNGSVSRVPAALQELRKNGVQIIRSSRQGGFVLRNAEQPDDKNDW 301
 Db 255 INAGVQGVASGVNGVNAALDLERKADSDVAVVARSVPTGYTTRDAEVDSDKGYF 314
 Qy 302 VVAHDLPQKARILAMVAMTKTQDSKEIQRIFWEX 336
 Db 315 VASGTLNPQKARVILQALTKTQDKPKVIOQIFEDP 349

RESULT 9

Asparaginase (EC 3.5.1.1) yccc [similarity] - *Bacillus subtilis*

C:Species: *Bacillus subtilis*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: F69754
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertorello, C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Erttington, J.; Fabret, C.; Ferrari, B.

Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galizzi, A.; Gallert, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Hullo, M.F.; Koester, P.; Koningsstein, G.; Krog, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols, A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogihara, A.; Oudgaa, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadleir, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Seror, akuchl, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumeitein, B.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:96044033; PMID:938377
 A:Accession: F69754
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-375 <KUN>
 A:Cross-references: UNIPROT:O34482; GB:Z99105; GB:AL009126; NID:g2632457; PIND:CAB12063.1
 C:Experimental source: strain 168
 C:Genetics:

A:Gene: yccC
C:Superfamily: asparaginase
C:Keywords: hydrolase

Query Match 45.1%; Score 758; DB 2; Length 375;
Best Local Similarity 47.6%; Pred. No. 2.9e-44;
Matches 158; Conservative 58; Mismatches 112; Indels 4; Gaps 3;

QY 6 OOKLNVVILATGCTTAGAGASAANSATYQAAKVGVDKLIAGPELADIANVREGVQVMI 65
DB 47 KOLPVRILATGCTTAGAGDQSTSTTKAGVGVESILEAVPEKDIANVSGEQIVAV 106
QY 66 ASESTINDDLKLASSVAELADSNVDGIVTHGDTLEETAYFLNLVEKTDKPIYVVS 125
DB 107 GSTINDKLLKLAKRINLLASDDVDGIVTHGDTLEETAYFLNLVYSKDEPIYVS 166
QY 126 MRPSTASADGMLNLYNAVAASNKDSRGKGLVTNDEIQSGRDVSKSINIKTEAFKS- 184
DB 167 MRPSTASADGPNLYNAVAAGAPAKGKGLVVLNDIASRKYVTKTNTTTTDTFKSE 226
QY 165 AMGPVGVESYMPRLPAKHVTNVEPDIKQISLPQVDIAYSXGNTDTAYKALAQN 244
DB 227 EMGFVGTIAD-DYFNNKILTRKHTKDTDFSVSNLDELPOVDILYQNDGSYLPDAVVA 285
QY 245 GAKALIHAGTNGSVSSRVVPAQLQELRKNGVQIIRSSRQGGFVLNNAEQPDKNDWVVA 304
DB 266 GAKGIVFAGSGNSLSDAEKADSAVKGIVVRETRTGNGVVTNQGAYR-KDLAS 343
QY 305 HDLNPQAKRILAMVAMTKTQDSKEIQRIFWEX 336
DB 344 NSLNPQAKRILMLALTKTNDPOKIQAYFWEX 375

RESULT 10

A:Accession: ADO169
A:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #ext_change 09-Jul-2004
C:Accession: ADO169
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmons, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: ADO169
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-345 <KUR->
A:Cross-references: UNIPROT:Q8ZGB7; GB:AL590842; PDB:CA090215.1; PID:g15979435; GSPDB:C
A:Gene: ansp
C:Superfamily: asparaginase
C:Keywords: hydrolase

Query Match 44.6%; Score 749.5; DB 2; Length 345;
Best Local Similarity 47.7%; Pred. No. 9.8e-44;
Matches 157; Conservative 54; Mismatches 111; Indels 7; Gaps 4;

QY 9 LANVVLATGCTTAGAGASAANSATYQAAKVGVDKLIAGPELADIANVREGVQVMI 68
DB 23 LFNITLILATGCTTAGAGDSATKS-NTTAKGLGVDLVEAVPAIKDIANVREGVQVMI 81
QY 69 SITNDLKLASSVAELADSNVDGIVTHGDTLEETAYFLNLVEKTDKPIYVVSMP 128
DB 82 DNNDVWMLTAKKINK--DCTKTDFVITHGDTLEETAYFLNLVYCNCKPVIYVGNMP 139
QY 129 GTAMASADGMLNLYNAVAASNKDSRGKGLVTNDEIQSGRDVSKSINIKTEAFKS-AW 187
DB 140 ATALGADGPLNLYNAVAASDADSAKRGVLANNDVFTGRDVKNTNTTSVQTPQSPNTG 199
QY 188 PLGMVGVESYMPRLPAKHVTNVEPDIKQISLPQVDIAYSXGNTDTAYKALAQNGAK 247

DB 200 PLGYIYDGKVIYIHOAPAQOLA---FDISKMLNTPRVGIYVYANVASDIPAKALADGYQ 256
QY 246 ALIHAGTNGSVSSRVVPAQLQELRKNGVQIIRSSRQGGFVLNNAEQPDKNDWVVAHD 307
DB 257 GIVSAGVGNGLYHTVPDTLATPAASHGAVAVSSRSPSTTEGABIDDAKYGFAAGAL 316
QY 308 NPQAKRILAMVAMTKTQDSKEIQRIFWEX 336
DB 317 NPQAKRILMLALTKTNDPOKIQAYFWEX 375

RESULT 11

A:Accession: A26054
A:Species: Brwnia chrysanthemi
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #ext_change 09-Jul-2004
C:Accession: A26054; S03681
R:Minon, N.P.; Bullman, H.M.S.; Scawen, M.D.; Atkinson, T.; Gilbert, H.J.
Gene 46, 25-35, 1986
A:Title: Nucleotide sequence of the Brwnia chrysanthemi NCPPB 1066 L-asparaginase gene.
A:Reference number: A26054; MUID:87106840; PMID:3026924
A:Accession: A26054
A:Molecule type: DNA
A:Residues: 1-348 <MIN->
A:Cross-references: UNIPROT:P06008; GB:M14741; GB:X14777; NID:942965; PIDN:CAA32884.1; P1
A:Note: The authors translated the codon AAG for residue 286 as Leu
R:Filipula, D.; Nagle, J.W.; Pulford, S.; Anderson, D.M.
Nucleic Acids Res. 16, 10385, 1988
A:Title: Sequence of L-asparaginase gene from Brwnia chrysanthemi NCPPB 1125.
A:Reference number: S03681; MUID:89057497; PMID:3194219
A:Accession: S03681
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-176, 'T', 178-198, 'R', 200-287, 'U', 289-294, 'W', 296-348 <FIL->
A:Cross-references: GB:X12746; NID:940993; PIDN:CAA31239.1; PID:g40994
C:Superfamily: asparaginase
C:Keywords: hydrolase
P:1-21/Domain: signal sequence #status predicted <SIG>
P:22-348/Product: asparaginase #status predicted <MAR>

Query Match 44.6%; Score 749; DB 1; Length 348;
Best Local Similarity 47.3%; Pred. No. 1.1e-43;
Matches 156; Conservative 63; Mismatches 105; Indels 6; Gaps 2;

QY 8 KLANVVLATGCTTAGAGASAANSATYQAAKVGVDKLIAGPELADIANVREGVQVMI 67
DB 24 LFNITLILATGCTTAGAGATYQTTGYKAGLGVDTLNAVEPVKLANVKGQPSNMAS 83
QY 68 SITNDLKLASSVAELADSNVDGIVTHGDTLEETAYFLNLVEKTDKPIYVVSMP 127
DB 84 ENMTGVDVILKSORVVELIARDVDGIVTHGDTLEETAYFLNLVYSKDEPIYVSMP 143
QY 128 PGTAMASADGMLNLYNAVAASNKDSRGKGLVTNDEIQSGRDVSKSINIKTEAFKS-SAW 186
DB 144 PATAISADGPMVLLEAVRVAAGDKQSRGCVVVLNDRIGASAYITKTNAJLTDTPKANE 203
QY 187 GELGVGVESYMPRLPAKHVTNVEPDIKQISLPQVDIAYSXGNTDTAYKALAQNGA 246
DB 204 GYLGVIIGNRIYQNRIDKLTTRSVFVGRGLTSPKVDILGYDDPEYLDAAIQGV 263
QY 247 KALIHAGTNGSVSSRVVPAQLQELRKNGVQIIRSSRQGGFVLNNAEQPDKNDWVVAHD 306
DB 264 KGIIVAGMGAGSVSRGILAGMKAMEKGVVIRSTRNGIV-----PDDELPLGLVSDS 318
QY 307 IMPQAKRILAMVAMTKTQDSKEIQRIFWEX 336
DB 319 IMPQAKRILMLALTKTNDPOKIQAYFWEX 375

RESULT 12

H81418
asparaginase (EC 3.5.1.1) Cytoplasmic Cj0029 [imported] - Campylobacter jejuni (strain NC
C:Species: Campylobacter jejuni

[illegible][illegible]

Db 301 DNLPQKARVLLQLALTKTNDKAKIQEMFEY 332

RESULT 15

T50284

aspartarginase (EC 3.5.1.1) precursor [similarity] - fission yeast (*Schizosaccharomyces pombe*)
C/Species: *Schizosaccharomyces pombe*

C:\Date: 09-Jun-2000 #sequence_rev181

C/Accession: T50284

R, Zimmermann, W.; W

submitted to the EMBL Data Library, January 2000

A/Reference number: Z25053

A: Accession: T50284

A, Status: preliminary

A/Molecule type: DNA

A/Residues: 1-356 <ZIM>

A;CROSS-references: INT

A/Cross-references: UNIPROT:Q9UTS7; EMBL:AL137130; PDB: CAB69634.1; GSPDB:GN00066; SPDB: A/Bioinformatics source: study 672b/-; PMID 2077

A/Experimental source: strain 972h(-); cosmid c977

C, Genetics:

A1Gene: SPDB:SPAC977.12

A/Mar position: 1
C/Superfamily: asparaginase

C1 Keywords: hydrolase

32.7%; Score 550; DB 2; Length 356;

Best Local Similarity 41.0%; Pred. No. 4e-30;

Matches 136; Conservative 56; Mismatches 120; Indels 20; Gaps 8;

9 LANVILATGCTIAGAGSAANSATYQAAKVGVDKLIAGVPELADLANVRGEQVMQIASE 68

Db 35 LPNTVIFAMGGTIAGCANSLEIWNYPGSGVIEKLI EAVPAIKAIANINGVQVTNMGSE 94

```
69 BLINDLEKLASSV-ARLAUSDNDGIVITHTGTLIBETAYFNLNVEKDKXPIWVGSMR 127
```

Db 95 NLT PADV LKAKLILAEVAKPN - VHGI VITHGTDSLEETAMFLDTISTAKPIVVVGMR 153

[illegible]

134 P81A1GADGPNULNAAVAVASSNQSMRGILVLENDRISSAFIILKINGNLDIFNSIEA 213

Db 214 GSLGIVLNQKPEYFFSPAV-PTGKVFPIYNIKQLPRVDILYNGGINPKLAESAVHIG 272

347 KAT HACTHCNCNVS---CQIUBA'OP1 BNACTIOTBCEBBOOCCCT BNAEBOBNNAM--W 301

Db 273 KGLVLAAMGATSWTDGNEVISSL-IREHNI PVVYSHRTABGY-----SSNSCLG 3211

302 VVAHDLPKARIILAMVAMTKTODSKEIORIF 333

Db 322 IPSYFLNPQKARYMLLAISSGYSIRDIEDLF 353

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Search completed: March 18, 2005, 21:33:21
Job time : 43 secs

Job time : 43 secs

This Page Blank (uspto)